GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 2, 2004, 16:15:45 ; Search time 2716 Seconds
(without alignments)
150.625 Million cell updates/sec

Title: US-09-875-453B-5
Perfect score: 10
Sequence: 1 gagttttgtt 10

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database :
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Search Notes

" Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	10	100.0	10	6	AX351053	AX351053 Sequence
c	2	10	100.0	21	6	AX537678	AX537678 Sequence
	3	10	100.0	36	6	A78758	A78758 Sequence 19
	4	10	100.0	36	6	AR014731	AR014731 Sequence
c	5	10	100.0	39	6	AR142981	AR142981 Sequence
c	6	10	100.0	47	6	AR288947	AR288947 Sequence
	7	10	100.0	49	8	ATH521017	AJ521017 Arabidops
c	8	10	100.0	50	6	AX162487	AX162487 Sequence
c	9	10	100.0	51	6	AX162488	AX162488 Sequence
	10	10	100.0	63	9	AY152466	AY152466 Homo sapi
c	11	10	100.0	65	6	AX483134	AX483134 Sequence
c	12	10	100.0	65	6	AX483267	AX483267 Sequence
c	13	10	100.0	65	6	AX484145	AX484145 Sequence
	14	10	100.0	65	6	AX485358	AX485358 Sequence
	15	10	100.0	65	6	AX485361	AX485361 Sequence
	16	10	100.0	65	6	AX486487	AX486487 Sequence
	17	10	100.0	70	17	HSA247013	Aj247013 Homo sapi
c	18	10	100.0	80	8	ATH527221	AJ527221 Arabidops
	19	10	100.0	97	4	SHPMCRE	L35313 Ovis aries
c	20	10	100.0	103	9	AF502326	AF502326 Macaca mu
c	21	10	100.0	103	9	AF502361	AF502361 Macaca mu
	22	10	100.0	117	8	ATH531685	AJ531685 Arabidops
	23	10	100.0	117	8	ATH531802	AJ531802 Arabidops
	24	10	100.0	117	8	ATH531828	AJ531828 Arabidops
c	25	10	100.0	117	8	ATH552575	AJ552575 Arabidops
	26	10	100.0	121	6	AX266705	AX266705 Sequence
c	27	10	100.0	121	6	AX266706	AX266706 Sequence
	28	10	100.0	125	6	AX196477	AX196477 Sequence
c	29	10	100.0	128	6	AX694812	AX694812 Sequence
c	30	10	100.0	129	8	ATH532306	AJ532306 Arabidops
	31	10	100.0	131	14	RBFMVSFVA	M22117 Malignant r
c	32	10	100.0	133	6	BD050001	BD050001 Sequence
c	33	10	100.0	135	11	HSPF15C4	Z94629 H.sapiens f
c	34	10	100.0	139	6	AX196355	AX196355 Sequence
	35	10	100.0	139	9	HS152H1F	Z59406 H.sapiens C
c	36	10	100.0	139	11	HSPA33A6	Z94241 H.sapiens f
	37	10	100.0	142	5	CLU55846	U55846 Columba liv
	38	10	100.0	142	6	BD053341	BD053341 Sequence
	39	10	100.0	144	6	E50611	E50611 Betacelluli
c	40	10	100.0	144	11	G26244	G26244 human STS T
	41	10	100.0	146	3	AG3UTRA41	Y08165 A.gambiae m
	42	10	100.0	147	6	E50610	E50610 Betacelluli
c	43	10	100.0	149	5	AF462550	AF462550 Hemibagru
	44	10	100.0	149	6	AX341152	AX341152 Sequence
	45	10	100.0	149	6	BD056116	BD056116 Sequence
	46	10	100.0	153	6	BD038248	BD038248 Sequence
	47	10	100.0	154	8	AY033692	AY033692 Arabidops
	48	10	100.0	154	11	G48354	G48354 SHGC-68482
c	49	10	100.0	155	6	AX072015	AX072015 Sequence
c	50	10	100.0	155	6	AX111502	AX111502 Sequence
	51	10	100.0	155	6	BD113958	BD113958 EST and e
	52	10	100.0	158	11	BX296010	BX296010 Arabidops
	53	10	100.0	159	8	AY202777	AY202777 Arabidops
	54	10	100.0	161	8	ATH524027	AJ524027 Arabidops
c	55	10	100.0	161	8	ATH526971	AJ526971 Arabidops
c	56	10	100.0	161	8	ATH526987	AJ526987 Arabidops
	57	10	100.0	162	6	A48603	A48603 Sequence 19
	58	10	100.0	163	11	G43942	G43942 WIAF-3564-S
	59	10	100.0	163	11	G59698	G59698 SHGC-130276
	60	10	100.0	164	8	ATH521661	AJ521661 Arabidops
	61	10	100.0	165	8	AY201583	AY201583 Arabidops
	62	10	100.0	165	8	AY201584	AY201584 Arabidops
c	63	10	100.0	166	8	ATH552282	AJ552282 Arabidops
c	64	10	100.0	168	11	AL834544	AL834544 Arabidops
c	65	10	100.0	169	11	AL834563	AL834563 Arabidops

c 66	10	100.0	173	6	BD039447	BD039447 Sequence
c 67	10	100.0	175	6	BD040263	BD040263 Sequence
c 68	10	100.0	177	5	AY212504	AY212504 Diodon hy
69	10	100.0	178	8	ATH505719	AJ505719 Arabidops
c 70	10	100.0	178	9	HSTRGG1R	X02757 H.sapiens g
71	10	100.0	179	6	BD039560	BD039560 Sequence
c 72	10	100.0	182	11	HSPE33A10	AL009809 H.sapiens
73	10	100.0	183	6	AX312694	AX312694 Sequence
c 74	10	100.0	186	11	BX294626	BX294626 Arabidops
c 75	10	100.0	188	11	HSPE05F05	AL009322 H.sapiens
76	10	100.0	189	3	AY143621	AY143621 Amphiphol
77	10	100.0	190	9	HS48E5F	Z61237 H.sapiens C
c 78	10	100.0	191	6	BD112830	BD112830 EST and e
79	10	100.0	191	11	HUMSWX1667	L41763 Human chrom
80	10	100.0	193	6	AX361556	AX361556 Sequence
81	10	100.0	195	11	BX296014	BX296014 Arabidops
c 82	10	100.0	198	14	AF280494	AF280494 Respirato
c 83	10	100.0	198	14	AF280495	AF280495 Respirato
c 84	10	100.0	198	14	AF280496	AF280496 Respirato
c 85	10	100.0	198	14	AF280497	AF280497 Respirato
c 86	10	100.0	198	14	AF280498	AF280498 Respirato
c 87	10	100.0	198	14	AF280499	AF280499 Respirato
c 88	10	100.0	198	14	AF280500	AF280500 Respirato
c 89	10	100.0	198	14	AF280501	AF280501 Respirato
c 90	10	100.0	198	14	AF280502	AF280502 Respirato
c 91	10	100.0	198	14	AF280503	AF280503 Respirato
c 92	10	100.0	198	14	AF280504	AF280504 Respirato
c 93	10	100.0	198	14	AF280505	AF280505 Respirato
c 94	10	100.0	198	14	AF280506	AF280506 Respirato
c 95	10	100.0	198	14	AF280507	AF280507 Respirato
c 96	10	100.0	198	14	AF280508	AF280508 Respirato
c 97	10	100.0	198	14	AF280509	AF280509 Respirato
c 98	10	100.0	198	14	AF280510	AF280510 Respirato
c 99	10	100.0	198	14	AF280511	AF280511 Respirato
c 100	10	100.0	198	14	AF280512	AF280512 Respirato
c 101	10	100.0	198	14	AF280513	AF280513 Respirato
c 102	10	100.0	198	14	AF280514	AF280514 Respirato
c 103	10	100.0	198	14	AF280515	AF280515 Respirato
c 104	10	100.0	198	14	AF280516	AF280516 Respirato
c 105	10	100.0	198	14	AF280517	AF280517 Respirato
c 106	10	100.0	198	14	AF280518	AF280518 Respirato
c 107	10	100.0	198	14	AF280519	AF280519 Respirato
c 108	10	100.0	198	14	AF280520	AF280520 Respirato
c 109	10	100.0	198	14	AF280521	AF280521 Respirato
c 110	10	100.0	198	14	AF280522	AF280522 Respirato
c 111	10	100.0	198	14	AF280523	AF280523 Respirato
c 112	10	100.0	198	14	AF280524	AF280524 Respirato
c 113	10	100.0	198	14	AF280525	AF280525 Respirato
c 114	10	100.0	198	14	AF280526	AF280526 Respirato
c 115	10	100.0	198	14	AF280527	AF280527 Respirato
c 116	10	100.0	198	14	AF280528	AF280528 Respirato
c 117	10	100.0	198	14	AF280529	AF280529 Respirato
c 118	10	100.0	198	14	AF280530	AF280530 Respirato
c 119	10	100.0	198	14	AF280531	AF280531 Respirato
c 120	10	100.0	198	14	AF280532	AF280532 Respirato
c 121	10	100.0	198	14	AF280533	AF280533 Respirato
c 122	10	100.0	198	14	AF280534	AF280534 Respirato
c 123	10	100.0	198	14	AF280535	AF280535 Respirato
c 124	10	100.0	198	14	AF280536	AF280536 Respirato
c 125	10	100.0	198	14	AF280537	AF280537 Respirato
c 126	10	100.0	198	14	AF280538	AF280538 Respirato
127	10	100.0	201	6	BD114177	BD114177 EST and e
c 128	10	100.0	202	6	A74736	A74736 Sequence 42
c 129	10	100.0	202	6	A77715	A77715 Sequence 42
c 130	10	100.0	204	6	AX390584	AX390584 Sequence
131	10	100.0	205	1	AF280074	AF280074 Propionib
c 132	10	100.0	205	11	DM56C8S	Z50606 D. melanoga
c 133	10	100.0	206	8	PSYGRAITS1	AF072054 Psychotri
c 134	10	100.0	208	11	BX296441	BX296441 Arabidops
135	10	100.0	213	8	BT004658	BT004658 Arabidops
c 136	10	100.0	213	11	G04407	G04407 human STS W
137	10	100.0	214	6	BD033361	BD033361 Sequence
c 138	10	100.0	214	6	BD050146	BD050146 Sequence
c 139	10	100.0	214	6	BD059583	BD059583 Secreted
140	10	100.0	214	11	G29800	G29800 human STS S
141	10	100.0	220	6	BD118467	BD118467 EST and e
142	10	100.0	221	9	HS56A3F	Z61642 H.sapiens C
143	10	100.0	222	6	A97176	A97176 Sequence 22
144	10	100.0	222	6	A97258	A97258 Sequence 10
145	10	100.0	222	6	BD076553	BD076553 Method of
146	10	100.0	222	6	BD076635	BD076635 Method of
147	10	100.0	224	5	AF108246	AF108246 Mugil cep
c 148	10	100.0	224	8	AY022619	AY022619 Oryza sat
149	10	100.0	224	9	HS197C3F	Z57794 H.sapiens C
150	10	100.0	225	8	AF073673S2	AF073674 Chaerophy

ALIGNMENTS

```
RESULT 1
AX351053
LOCUS       AX351053                  10 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 5 from Patent WO0194600.
ACCESSION   AX351053
VERSION     AX351053.1  GI:18616407
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kim,J.P., Starr,D.B., Tam,A.W., Laurance,M.E., Michelotti,E.F.,
            Velligan,M.D., Latour,D.R., Thomas,R.L., Kongpachith,A.,
            Sheppard,L.T., Lim,M.Y. and Bruice,T.W.
  TITLE     Promoters for regulated gene expression
  JOURNAL   Patent: WO 0194600-A 5 13-DEC-2001;
            GENELABS TECHNOLOGIES, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .10
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT        1 a      0 c      3 g      6 t
ORIGIN

  Query Match           100.0%;  Score 10;  DB 6;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          1 GAGTTTTGTT 10


RESULT 2
AX537678/c
LOCUS       AX537678                  21 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 28 from Patent EP1241269.
ACCESSION   AX537678
VERSION     AX537678.1  GI:25269647
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Heiskala,M.
  TITLE     Method for detecting reg-like protein and nucleic acids coding
            therefor
  JOURNAL   Patent: EP 1241269-A 28 18-SEP-2002;
            Ortho-Clinical Diagnostics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
```

```
                     /note="Artificial"
BASE COUNT        9 a      7 c      1 g      4 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         21 GAGTTTTGTT 12


RESULT 3
A78758
LOCUS       A78758                    36 bp    DNA     linear   PAT 19-OCT-1999
DEFINITION  Sequence 19 from Patent EP0566525.
ACCESSION   A78758
VERSION     A78758.1  GI:6090360
KEYWORDS    .
SOURCE      Impatiens necrotic spot virus
  ORGANISM  Impatiens necrotic spot virus
            Viruses; ssRNA negative-strand viruses; Bunyaviridae; Tospovirus.
REFERENCE   1  (bases 1 to 36)
  AUTHORS   Gielen,J.J. and Goldbach,R.W.
  TITLE     RECOMBINANT TOSPOVIRUS DNA CONSTRUCTS AND PLANTS COMPRISING SUCH
            CONSTRUCTS
  JOURNAL   Patent: EP 0566525-A 19 20-OCT-1993;
            SANDOZ LTD (CH); SANDOZ AG (DE)
FEATURES             Location/Qualifiers
     source          1. .36
                     /organism="Impatiens necrotic spot virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11612"
BASE COUNT        5 a      5 c      9 g     17 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 36;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         13 GAGTTTTGTT 22


RESULT 4
AR014731
LOCUS       AR014731                  36 bp    DNA     linear   PAT 05-DEC-1998
DEFINITION  Sequence 19 from patent US 5773700.
ACCESSION   AR014731
VERSION     AR014731.1  GI:3972185
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 36)
  AUTHORS   Van Grinsven,M.Quirinius.Joseph.Marie., De Haan,P.Theodorus.,
            Gielen,J.Jacobus.Ludgerus., Peters,D. and Goldbach,R.Willem.
  TITLE     Constructs containing impatiens necrotic spot tospovirus RNA and
            methods of use thereof
  JOURNAL   Patent: US 5773700-A 19 30-JUN-1998;
FEATURES             Location/Qualifiers
     source          1. .36
                     /organism="unknown"
BASE COUNT        5 a      5 c      9 g     17 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 36;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         13 GAGTTTTGTT 22


RESULT 5
AR142981/c
LOCUS       AR142981                  39 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 6 from patent US 6204035.
ACCESSION   AR142981
VERSION     AR142981.1  GI:15104267
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Wiedmer,T. and Sims,P.J.
  TITLE     Methods and compositions to alter the cell surface expression of
            phosphatidylserine and other clot-promoting plasma membrane
            phospholipids
  JOURNAL   Patent: US 6204035-A 6 20-MAR-2001;
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="unknown"
BASE COUNT       16 a      9 c      7 g      7 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         32 GAGTTTTGTT 23


RESULT 6
AR288947/c
LOCUS       AR288947                  47 bp    DNA     linear   PAT 12-JUN-2003
DEFINITION  Sequence 682 from patent US 6537751.
ACCESSION   AR288947
VERSION     AR288947.1  GI:31676231
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Cohen,D., Chumakov,I. and Blumenfeld,M.
  TITLE     Biallelic markers for use in constructing a high density
            disequilibrium map of the human genome
  JOURNAL   Patent: US 6537751-A 682 25-MAR-2003;
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="unknown"
BASE COUNT       22 a      5 c      7 g     12 t      1 others
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 47;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         21 GAGTTTTGTT 12


RESULT 7
ATH521017
LOCUS       ATH521017                 49 bp    DNA     linear   PLN 29-MAR-2003
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence, left border, clone
            050E05.
ACCESSION   AJ521017
VERSION     AJ521017.1  GI:26789253
KEYWORDS    left border; T-DNA flanking sequence.
```

```
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
            Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
            Lepiniec,L., Caboche,M. and Lecharny,A.
  TITLE     T-DNA integration into the Arabidopsis genome depends on sequences
            of pre-insertion sites
  JOURNAL   EMBO Rep. 3 (12), 1152-1157 (2002)
  MEDLINE   22363535
   PUBMED   12446565
REFERENCE   2  (bases 1 to 49)
  AUTHORS   Balzergue,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
            Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT     PCR was performed on DNA from transformants of Arabidopsis thaliana
            plants from INRA (Versailles). The DNA fragment(s) resulting from
            the PCR were directly sequenced from the left or the right border
            to determine the genomic sequence flanking the insertion. T-DNA
            derived sequences were removed. Information to order the
            corresponding mutant line and a link to a database providing a
            graphical display of the insertion site are available at
            http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
            been generated in the framework of the French plant genomics
            program 'Genoplante' (http://www.genoplante.com and
            http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /cultivar="Wassillewskija"
                     /db_xref="taxon:3702"
                     /clone="050E05"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
     misc_feature    1. .49
                     /note="T-DNA flanking sequence
                     left border"
BASE COUNT        7 a      7 c     11 g     24 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 8;  Length 49;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         16 GAGTTTTGTT 25


RESULT 8
AX162487/c
LOCUS       AX162487                  50 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 5815 from Patent WO0140521.
ACCESSION   AX162487
VERSION     AX162487.1  GI:14543818
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 5815 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     misc_feature    25. .26
                     /note="Nucleotide deleted between bases 25 and 26
                     Accession number cg44036050"
     misc_feature    26
                     /note="1 of 2 allelic variants (5816 is other entry)"
BASE COUNT       24 a     11 c      8 g      7 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         20 GAGTTTTGTT 11


RESULT 9
AX162488/c
LOCUS       AX162488                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 5816 from Patent WO0140521.
ACCESSION   AX162488
VERSION     AX162488.1  GI:14543819
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 5816 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (5815 is other entry)
                     Accession number cg44036050"
BASE COUNT       25 a     11 c      8 g      7 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         20 GAGTTTTGTT 11


RESULT 10
AY152466
LOCUS       AY152466                  63 bp    DNA     linear   PRI 15-DEC-2002
DEFINITION  Homo sapiens isolate 17 RUNX1/CBFA2T1 translocation breakpoint
            sequence.
ACCESSION   AY152466
VERSION     AY152466.1  GI:26984023
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   Zhang,Y., Strissel,P., Strick,R., Chen,J., Nucifora,G., Le
            Beau,M.M., Larson,R.A. and Rowley,J.D.
  TITLE     Genomic DNA breakpoints in AML1/RUNX1 and ETO cluster with
            topoisomerase II DNA cleavage and DNase I hypersensitive sites in
```

```
          t(8;21) leukemia
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (5), 3070-3075 (2002)
  MEDLINE   21874099
   PUBMED   11867721
REFERENCE   2  (bases 1 to 63)
  AUTHORS   Zhang,Y. and Rowley,J.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-SEP-2002) Department of Medicine, University of
          Chicago, 5841 S. Maryland Ave., MC2115, Chicago, IL 60637, USA
FEATURES             Location/Qualifiers
     source          1. .63
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolate="17"
                     /isolation_source="acute myeloid leukemia patient"
                     /db_xref="taxon:9606"
                     /map="t(21;8) (q22;q22)"
     source          1. .14
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolation_source="acute myeloid leukemia patient"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22"
     source          33. .63
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /isolation_source="acute myeloid leukemia patient"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8q22"
     misc_feature    1. .63
                     /note="RUNX1/CBFA2T1 translocation breakpoint region"
     gene            <1. .14
                     /gene="RUNX1"
                     /note="synonym: AML1"
     intron          <1. .14
                     /gene="RUNX1"
                     /number=5
     misc_feature    15. .32
                     /note="nontemplated insertion of unknown origin"
     gene            33. .>63
                     /gene="CBFA2T1"
                     /note="synonym: ETO"
     intron          33. .>63
                     /gene="CBFA2T1"
                     /number=1b
BASE COUNT       16 a     11 c     12 g     24 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 9;  Length 63;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         47 GAGTTTTGTT 56


RESULT 11
AX483134/c
LOCUS       AX483134                 65 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 434 from Patent WO02053728.
ACCESSION   AX483134
VERSION     AX483134.1  GI:22317554
KEYWORDS    .
SOURCE      Candida albicans
  ORGANISM  Candida albicans
          Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
          Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 434 11-JUL-2002;
          Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Candida albicans"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:5476"
BASE COUNT       22 a     20 c      3 g     20 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 65;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         39 GAGTTTTGTT 30


RESULT 12
AX483267/c
LOCUS       AX483267                 65 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 567 from Patent WO02053728.
ACCESSION   AX483267
VERSION     AX483267.1  GI:22317687
KEYWORDS    .
SOURCE      Candida albicans
  ORGANISM  Candida albicans
          Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
          Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 567 11-JUL-2002;
          Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Candida albicans"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:5476"
BASE COUNT       27 a     12 c      4 g     22 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 65;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         16 GAGTTTTGTT 7


RESULT 13
AX484145/c
LOCUS       AX484145                 65 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 1445 from Patent WO02053728.
ACCESSION   AX484145
VERSION     AX484145.1  GI:22318497
KEYWORDS    .
SOURCE      Candida albicans
  ORGANISM  Candida albicans
          Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
          Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 1445 11-JUL-2002;
          Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Candida albicans"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:5476"
```

```
BASE COUNT       32 a     13 c     10 g     10 t
ORIGIN

  Query Match            100.0%;  Score 10;  DB 6;  Length 65;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTGTT 10
              ||||||||||
Db         35 GAGTTTGTT 26


RESULT 14
AX485358
LOCUS       AX485358                  65 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 2658 from Patent WO02053728.
ACCESSION   AX485358
VERSION     AX485358.1  GI:22319642
KEYWORDS    .
SOURCE      Candida albicans
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 2658 11-JUL-2002;
            Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Candida albicans"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:5476"
BASE COUNT       15 a      8 c      9 g     33 t
ORIGIN

  Query Match            100.0%;  Score 10;  DB 6;  Length 65;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTGTT 10
              ||||||||||
Db         43 GAGTTTGTT 52


RESULT 15
AX485361
LOCUS       AX485361                  65 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 2661 from Patent WO02053728.
ACCESSION   AX485361
VERSION     AX485361.1  GI:22319645
KEYWORDS    .
SOURCE      Candida albicans
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 2661 11-JUL-2002;
            Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Candida albicans"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:5476"
BASE COUNT       22 a     11 c      8 g     24 t
ORIGIN

  Query Match            100.0%;  Score 10;  DB 6;  Length 65;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTGTT 10
              ||||||||||
Db         16 GAGTTTGTT 25


RESULT 16
AX486487
LOCUS       AX486487                  65 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 3787 from Patent WO02053728.
ACCESSION   AX486487
VERSION     AX486487.1  GI:22320703
KEYWORDS    .
SOURCE      Candida albicans
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 3787 11-JUL-2002;
            Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Candida albicans"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:5476"
BASE COUNT       19 a     13 c     10 g     23 t
ORIGIN

  Query Match            100.0%;  Score 10;  DB 6;  Length 65;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTGTT 10
              ||||||||||
Db         17 GAGTTTGTT 26


RESULT 17
HSA247013
ID   HSA247013  standard; DNA; HUM; 70 BP.
XX
AC   AJ247013;
XX
SV   AJ247013.1
XX
DT   24-JUN-1999 (Rel. 60, Created)
DT   24-JUN-1999 (Rel. 60, Last updated, Version 1)
XX
DE   Homo sapiens PAC trapped exon, clone 85M6 (70 bp)
XX
KW   PAC; trapped exon.
XX
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN   [1]
RP   1-70
RA   Niederfuehr A.;
RT   ;
RL   Submitted (22-JUN-1999) to the EMBL/GenBank/DDBJ databases.
RL   Niederfuehr A., Physiologische Chemie I, Theodor-Boveri-Institut fuer
RL   Biowissenschaften, am Hubland, D-97074 Wuerzburg, GERMANY.
XX
RN   [2]
RA   Niederfuehr A.;
RT   ;
RL   Thesis (1999), Universitaet Wuerzburg
XX
FH   Key             Location/Qualifiers
FH
FT   source          1. .70
```

```
FT                   /chromosome="11"
FT                   /db_xref="taxon:9606"
FT                   /mol_type="genomic DNA"
FT                   /organism="Homo sapiens"
FT                   /clone_lib="RPCI PAC 1,3-5"
FT                   /clone="85M6"
FT                   /map="11p13"
FT   exon            1. .70
FT                   /note="trapped"
XX
SQ   Sequence 70 BP; 19 A; 14 C; 16 G; 21 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 17;  Length 70;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         16 GAGTTTTGTT 25


RESULT 18
ATH527221/c
LOCUS       ATH527221                80 bp    DNA     linear   PLN 29-MAR-2003
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence, left border, clone
            135F01.
ACCESSION   AJ527221
VERSION     AJ527221.1  GI:26795481
KEYWORDS    left border; T-DNA flanking sequence.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
            Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
            Lepiniec,L., Caboche,M. and Lecharny,A.
  TITLE     T-DNA integration into the Arabidopsis genome depends on sequences
            of pre-insertion sites
  JOURNAL   EMBO Rep. 3 (12), 1152-1157 (2002)
  MEDLINE   22363535
   PUBMED   12446565
REFERENCE   2  (bases 1 to 80)
  AUTHORS   Balzergue,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
            Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT     PCR was performed on DNA from transformants of Arabidopsis thaliana
            plants from INRA (Versailles). The DNA fragment(s) resulting from
            the PCR were directly sequenced from the left or the right border
            to determine the genomic sequence flanking the insertion. T-DNA
            derived sequences were removed. Information to order the
            corresponding mutant line and a link to a database providing a
            graphical display of the insertion site are available at
            http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
            been generated in the framework of the French plant genomics
            program 'Genoplante' (http://www.genoplante.com and
            http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .80
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /cultivar="Wassillewskija"
                     /db_xref="taxon:3702"
                     /clone="135F01"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
     misc_feature    1. .80
                     /note="T-DNA flanking sequence
                     left border"
BASE COUNT       32 a     15 c      9 g     24 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 8;  Length 80;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         14 GAGTTTTGTT 5


RESULT 19
SHPMCRE
LOCUS       SHPMCRE                  97 bp    DNA     linear   MAM 16-AUG-1994
DEFINITION  Ovis aries DNA microsatellite.
ACCESSION   L35313
VERSION     L35313.1  GI:530198
KEYWORDS    PCR amplified; microsatellite; polymorphic microsatellite.
SOURCE      Ovis aries (sheep)
  ORGANISM  Ovis aries
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis.
REFERENCE   1  (bases 1 to 97)
  AUTHORS   Smith,A.J., Hulme,D.J. and Beh,K.J.
  TITLE     Five polymorphic ovine microsatellites
  JOURNAL   Unpublished (1994)
COMMENT     Original source text: Ovis aries DNA.
FEATURES             Location/Qualifiers
     source          1. .97
                     /organism="Ovis aries"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9940"
     primer_bind     complement(12. .37)
                     /note="binding site for PCR primer to amplify
                     microsatellite; putative"
     satellite       35. .75
                     /standard_name="perfect GT microsatellite"
                     /note="putative"
                     /rpt_type=other
     primer_bind     73. .96
                     /standard_name="binding site for reverse PCR primer"
                     /note="putative"
BASE COUNT       17 a     11 c     29 g     40 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 4;  Length 97;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          6 GAGTTTTGTT 15


RESULT 20
AF502326/c
LOCUS       AF502326                103 bp    DNA     linear   PRI 05-JUN-2002
DEFINITION  Macaca mulatta isolate C1 69 nuclear mitochochondrial gene
            sequence.
ACCESSION   AF502326
VERSION     AF502326.1  GI:21326062
KEYWORDS    .
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Vartanian,J.P. and Wain-Hobson,S.
  TITLE     Analysis of a library of macaque nuclear mitochondrial sequences
            confirms macaque origin of divergent sequences from old oral polio
            vaccine samples
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (11), 7566-7569 (2002)
  MEDLINE   22028984
```

```
   PUBMED   12032323
REFERENCE   2  (bases 1 to 103)
  AUTHORS   Vartanian,J.-P. and Wain-Hobson,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2002) Virology, Institut Pasteur, 28 rue du
            docteur Roux, Paris 75724, France
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Macaca mulatta"
                     /mol_type="genomic DNA"
                     /isolate="Cl 69"
                     /db_xref="taxon:9544"
     misc_feature    1. .103
                     /note="nuclear mitochondrial sequence; numts"
BASE COUNT       36 a     26 c     14 g     27 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 9;  Length 103;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         39 GAGTTTTGTT 30


RESULT 21
AF502361/c
LOCUS       AF502361                 103 bp    DNA     linear   PRI 05-JUN-2002
DEFINITION  Macaca mulatta isolate Cl 177 nuclear mitochochondrial gene
            sequence.
ACCESSION   AF502361
VERSION     AF502361.1  GI:21326097
KEYWORDS    .
SOURCE      Macaca mulatta (rhesus monkey)
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 103)
  AUTHORS   Vartanian,J.P. and Wain-Hobson,S.
  TITLE     Analysis of a library of macaque nuclear mitochondrial sequences
            confirms macaque origin of divergent sequences from old oral polio
            vaccine samples
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (11), 7566-7569 (2002)
  MEDLINE   22028984
   PUBMED   12032323
REFERENCE   2  (bases 1 to 103)
  AUTHORS   Vartanian,J.-P. and Wain-Hobson,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-APR-2002) Virology, Institut Pasteur, 28 rue du
            docteur Roux, Paris 75724, France
FEATURES             Location/Qualifiers
     source          1. .103
                     /organism="Macaca mulatta"
                     /mol_type="genomic DNA"
                     /isolate="Cl 177"
                     /db_xref="taxon:9544"
     misc_feature    1. .103
                     /note="nuclear mitochondrial sequence; numts"
BASE COUNT       35 a     27 c     14 g     27 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 9;  Length 103;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         39 GAGTTTTGTT 30


RESULT 22
ATH531685
LOCUS       ATH531685                117 bp    DNA     linear   PLN 29-MAR-2003
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence, left border, clone
            240B08.
ACCESSION   AJ531685
VERSION     AJ531685.1  GI:26799945
KEYWORDS    left border; T-DNA flanking sequence.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
            Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
            Lepiniec,L., Caboche,M. and Lecharny,A.
  TITLE     T-DNA integration into the Arabidopsis genome depends on sequences
            of pre-insertion sites
  JOURNAL   EMBO Rep. 3 (12), 1152-1157 (2002)
  MEDLINE   22363535
   PUBMED   12446565
REFERENCE   2  (bases 1 to 117)
  AUTHORS   Balzergue,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
            Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT     PCR was performed on DNA from transformants of Arabidopsis thaliana
            plants from INRA (Versailles). The DNA fragment(s) resulting from
            the PCR were directly sequenced from the left or the right border
            to determine the genomic sequence flanking the insertion. T-DNA
            derived sequences were removed. Information to order the
            corresponding mutant line and a link to a database providing a
            graphical display of the insertion site are available at
            http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
            been generated in the framework of the French plant genomics
            program 'Genoplante' (http://www.genoplante.com and
            http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .117
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /cultivar="Wassillewskija"
                     /db_xref="taxon:3702"
                     /clone="240B08"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
     misc_feature    1. .117
                     /note="T-DNA flanking sequence
                     left border"
BASE COUNT       38 a     17 c     24 g     38 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 8;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         95 GAGTTTTGTT 104


RESULT 23
ATH531802
LOCUS       ATH531802                117 bp    DNA     linear   PLN 29-MAR-2003
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence, left border, clone
            244A03.
ACCESSION   AJ531802
VERSION     AJ531802.1  GI:26800062
KEYWORDS    left border; T-DNA flanking sequence.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
```

```
REFERENCE   1
  AUTHORS   Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
            Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
            Lepiniec,L., Caboche,M. and Lecharny,A.
  TITLE     T-DNA integration into the Arabidopsis genome depends on sequences
            of pre-insertion sites
  JOURNAL   EMBO Rep. 3 (12), 1152-1157 (2002)
  MEDLINE   22363535
   PUBMED   12446565
REFERENCE   2  (bases 1 to 117)
  AUTHORS   Balzergue,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
            Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT     PCR was performed on DNA from transformants of Arabidopsis thaliana
            plants from INRA (Versailles). The DNA fragment(s) resulting from
            the PCR were directly sequenced from the left or the right border
            to determine the genomic sequence flanking the insertion. T-DNA
            derived sequences were removed. Information to order the
            corresponding mutant line and a link to a database providing a
            graphical display of the insertion site are available at
            http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
            been generated in the framework of the French plant genomics
            program 'Genoplante' (http://www.genoplante.com and
            http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .117
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /cultivar="Wassillewskija"
                     /db_xref="taxon:3702"
                     /clone="244A03"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
     misc_feature    1. .117
                     /note="T-DNA flanking sequence
                     left border"
BASE COUNT       38 a     17 c     24 g     38 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 8;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         95 GAGTTTTGTT 104


RESULT 24
ATH531828
LOCUS       ATH531828               117 bp    DNA     linear   PLN 29-MAR-2003
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence, left border, clone
            244G01.
ACCESSION   AJ531828
VERSION     AJ531828.1  GI:26800088
KEYWORDS    left border; T-DNA flanking sequence.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
            Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
            Lepiniec,L., Caboche,M. and Lecharny,A.
  TITLE     T-DNA integration into the Arabidopsis genome depends on sequences
            of pre-insertion sites
  JOURNAL   EMBO Rep. 3 (12), 1152-1157 (2002)
  MEDLINE   22363535
   PUBMED   12446565
REFERENCE   2  (bases 1 to 117)
  AUTHORS   Balzergue,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
            Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT     PCR was performed on DNA from transformants of Arabidopsis thaliana
            plants from INRA (Versailles). The DNA fragment(s) resulting from
            the PCR were directly sequenced from the left or the right border
            to determine the genomic sequence flanking the insertion. T-DNA
            derived sequences were removed. Information to order the
            corresponding mutant line and a link to a database providing a
            graphical display of the insertion site are available at
            http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
            been generated in the framework of the French plant genomics
            program 'Genoplante' (http://www.genoplante.com and
            http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .117
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /cultivar="Wassillewskija"
                     /db_xref="taxon:3702"
                     /clone="244G01"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
     misc_feature    1. .117
                     /note="T-DNA flanking sequence
                     left border"
BASE COUNT       38 a     17 c     24 g     38 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 8;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         95 GAGTTTTGTT 104


RESULT 25
ATH552575/c
LOCUS       ATH552575               117 bp    DNA     linear   PLN 29-MAR-2003
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence, left border, clone
            339A11.
ACCESSION   AJ552575
VERSION     AJ552575.1  GI:29368722
KEYWORDS    left border; T-DNA flanking sequence.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
            Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
            Lepiniec,L., Caboche,M. and Lecharny,A.
  TITLE     T-DNA integration into the Arabidopsis genome depends on sequences
            of pre-insertion sites
  JOURNAL   EMBO Rep. 3 (12), 1152-1157 (2002)
  MEDLINE   22363535
   PUBMED   12446565
REFERENCE   2  (bases 1 to 117)
  AUTHORS   Balzergue,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
            Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT     PCR was performed on DNA from transformants of Arabidopsis thaliana
            plants from INRA (Versailles). The DNA fragment(s) resulting from
            the PCR were directly sequenced from the left or the right border
            to determine the genomic sequence flanking the insertion. T-DNA
            derived sequences were removed. Information to order the
            corresponding mutant line and a link to a database providing a
            graphical display of the insertion site are available at
            http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
            been generated in the framework of the French plant genomics
            program 'Genoplante' (http://www.genoplante.com and
```

```
          http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .117
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /cultivar="Wassillewskija"
                     /db_xref="taxon:3702"
                     /clone="339A11"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
     misc_feature    1. .117
                     /note="T-DNA flanking sequence
                     left border"
BASE COUNT       43 a     11 c     13 g     50 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 8;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         92 GAGTTTTGTT 83


RESULT 26
AX266705
LOCUS       AX266705                121 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 4096 from Patent WO0173002.
ACCESSION   AX266705
VERSION     AX266705.1  GI:16515504
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kmiec,E.B., Gamper,H.B. and Rice,M.C.
  TITLE     Targeted chromosomal genomic alterations with modified single
            stranded oligonucleotides
  JOURNAL   Patent: WO 0173002-A 4096 04-OCT-2001;
            UNIVERSITY OF DELAWARE (US)
FEATURES             Location/Qualifiers
     source          1. .121
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT       23 a     25 c     31 g     42 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 121;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         93 GAGTTTTGTT 102


RESULT 27
AX266706/c
LOCUS       AX266706                121 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 4097 from Patent WO0173002.
ACCESSION   AX266706
VERSION     AX266706.1  GI:16515505
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kmiec,E.B., Gamper,H.B. and Rice,M.C.
  TITLE     Targeted chromosomal genomic alterations with modified single
            stranded oligonucleotides
  JOURNAL   Patent: WO 0173002-A 4097 04-OCT-2001;
            UNIVERSITY OF DELAWARE (US)
FEATURES             Location/Qualifiers
     source          1. .121
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT       42 a     31 c     25 g     23 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 121;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         29 GAGTTTTGTT 20


RESULT 28
AX196477
LOCUS       AX196477                125 bp    DNA     linear   PAT 07-SEP-2001
DEFINITION  Sequence 184 from Patent WO0151627.
ACCESSION   AX196477
VERSION     AX196477.1  GI:15386683
KEYWORDS    .
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1
  AUTHORS   Hauge,B.M., Wang,M.L., Parsons,J.D. and Parnell,L.D.
  TITLE     Nucleic acid molecules and other molecules associated with soybean
            cyst nematode resistance
  JOURNAL   Patent: WO 0151627-A 184 19-JUL-2001;
            MONSANTO COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .125
                     /organism="Glycine max"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3847"
                     /note="Seq ID: 318013_region_A3__93061_14"
BASE COUNT       45 a     22 c     23 g     35 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 125;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         35 GAGTTTTGTT 44


RESULT 29
AX694812/c
LOCUS       AX694812                128 bp    DNA     linear   PAT 31-MAR-2003
DEFINITION  Sequence 439 from Patent WO03008583.
ACCESSION   AX694812
VERSION     AX694812.1  GI:29417924
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Morris,D.W. and Engelhard,E.K.
  TITLE     Novel compositions and methods for cancer
  JOURNAL   Patent: WO 03008583-A 439 30-JAN-2003;
            Sagres Discovery (US)
FEATURES             Location/Qualifiers
```

```
     source          1. .128
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
BASE COUNT       57 a     22 c     29 g     20 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 128;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         33 GAGTTTTGTT 24


RESULT 30
ATH532306/c
LOCUS       ATH532306               129 bp    DNA     linear   PLN 29-MAR-2003
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence, left border, clone
            256G12.
ACCESSION   AJ532306
VERSION     AJ532306.1  GI:26800606
KEYWORDS    left border; T-DNA flanking sequence.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
            Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
            Lepiniec,L., Caboche,M. and Lecharny,A.
  TITLE     T-DNA integration into the Arabidopsis genome depends on sequences
            of pre-insertion sites
  JOURNAL   EMBO Rep. 3 (12), 1152-1157 (2002)
  MEDLINE   22363535
   PUBMED   12446565
REFERENCE   2  (bases 1 to 129)
  AUTHORS   Balzergue,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
            Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT     PCR was performed on DNA from transformants of Arabidopsis thaliana
            plants from INRA (Versailles). The DNA fragment(s) resulting from
            the PCR were directly sequenced from the left or the right border
            to determine the genomic sequence flanking the insertion. T-DNA
            derived sequences were removed. Information to order the
            corresponding mutant line and a link to a database providing a
            graphical display of the insertion site are available at
            http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
            been generated in the framework of the French plant genomics
            program 'Genoplante' (http://www.genoplante.com and
            http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .129
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /cultivar="Wassillewskija"
                     /db_xref="taxon:3702"
                     /clone="256G12"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
     misc_feature    1. .129
                     /note="T-DNA flanking sequence
                     left border"
BASE COUNT       32 a     29 c     23 g     45 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 8;  Length 129;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         96 GAGTTTTGTT 87


RESULT 31
RBFMVSFVA
LOCUS       RBFMVSFVA               131 bp    DNA     linear   VRL 03-AUG-1993
DEFINITION  Malignant rabbit fibroma virus Myxoma virus/Shope fibroma virus
            recombination site DNA.
ACCESSION   M22117
VERSION     M22117.1  GI:333602
KEYWORDS    recombination.
SOURCE      Malignant rabbit fibroma virus
  ORGANISM  Malignant rabbit fibroma virus
            Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
            Leporipoxvirus.
REFERENCE   1  (bases 1 to 131)
  AUTHORS   Upton,C., Macen,J.L., Maranchuk,R.A., DeLange,A.M. and McFadden,G.
  TITLE     Tumorigenic poxviruses: fine analysis of the recombination
            junctions in malignant rabbit fibroma virus, a recombinant between
            Shope fibroma virus and myxoma virus
  JOURNAL   Virology 166 (1), 229-239 (1988)
  MEDLINE   88322873
   PUBMED   2842947
COMMENT     Original source text: Malignant rabbit fibroma virus DNA.
FEATURES             Location/Qualifiers
     source          1. .131
                     /organism="Malignant rabbit fibroma virus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10274"
BASE COUNT       37 a     22 c     27 g     45 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 14;  Length 131;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         50 GAGTTTTGTT 59


RESULT 32
BD050001/c
LOCUS       BD050001                133 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Sequence tag and encoded human protein.
ACCESSION   BD050001
VERSION     BD050001.1  GI:22591743
KEYWORDS    JP 2001269182-A/26247.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 133)
  AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
  TITLE     Sequence tag and encoded human protein
  JOURNAL   Patent: JP 2001269182-A 26247 02-OCT-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001269182-A/26247
            PD   02-OCT-2001
            PF   24-FEB-2000 JP 2000118773
            PR   26-FEB-1999 US    60/122487
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
            PI   JORDAN
            PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
            PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
              G06F15/40
            CC
            FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .133
```

```
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT       36 a      32 c      38 g      27 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 133;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db        125 GAGTTTTGTT 116


RESULT 33
HSPF15C4/c
LOCUS       HSPF15C4                  135 bp    DNA     linear   STS 21-MAY-1998
DEFINITION  H.sapiens flow-sorted chromosome 20 HindIII fragment, SC20pF15C4,
            sequence tagged site.
ACCESSION   Z94629
VERSION     Z94629.1  GI:1946114
KEYWORDS    STS; single read.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 135)
  AUTHORS   Deloukas,P., Buck,D., Langford,C., Ross,M.T. and Hunt,S.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-APR-1997) The Sanger Centre, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
            humquery@sanger.ac.uk
COMMENT     Vector: pBSIISK+
            Marker stSG25547 (Primer A : TCAGCCTACACCTTGTTCCC; Primer B :
            GCAGCTCAAAAAGCAGATCC; amplimer size : 98 bp) was mapped to
            chromosome 20 using Radiation Hybrid panel Genebridge 4 (GB4).
FEATURES             Location/Qualifiers
     source          1. .135
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /clone="SC20pF15C4"
                     /sex="Female"
                     /tissue_type="EBV lymphoblastoid cell line"
                     /clone_lib="SC20pF"
                     /dev_stage="adult"
BASE COUNT       39 a      37 c      17 g      42 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 11;  Length 135;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         40 GAGTTTTGTT 31


RESULT 34
AX196355/c
LOCUS       AX196355                  139 bp    DNA     linear   PAT 07-SEP-2001
DEFINITION  Sequence 62 from Patent WO0151627.
ACCESSION   AX196355
VERSION     AX196355.1  GI:15386561
KEYWORDS    .
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.

REFERENCE   1
  AUTHORS   Hauge,B.M., Wang,M.L., Parsons,J.D. and Parnell,L.D.
  TITLE     Nucleic acid molecules and other molecules associated with soybean
            cyst nematode resistance
  JOURNAL   Patent: WO 0151627-A 62 19-JUL-2001;
            MONSANTO COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .139
                     /organism="Glycine max"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3847"
                     /note="Seq ID: 240017_region_G3__50537_17"
BASE COUNT       63 a      23 c      12 g      41 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 139;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         76 GAGTTTTGTT 67


RESULT 35
HS152H1F
LOCUS       HS152H1F                  139 bp    DNA     linear   PRI 19-OCT-1995
DEFINITION  H.sapiens CpG island DNA genomic Mse1 fragment, clone 152h1,
            forward read cpg152h1.ft1a.
ACCESSION   Z59406
VERSION     Z59406.1  GI:1031319
KEYWORDS    CpG island; genomic Mse1 fragment.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
  TITLE     Purification of CpG islands using a methylated DNA binding column
  JOURNAL   Nat. Genet. 6 (3), 236-244 (1994)
  MEDLINE   94282070
   PUBMED   8012384
REFERENCE   2  (bases 1 to 139)
  AUTHORS   Dodsworth,S.J., Huckle,E., Wilkinson,P. and Micklem,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk
COMMENT     Vector: pGEM-5Zf(-)
            Clones are available from the UK MRC Human Genome Mapping Project
            Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
            http://www.hgmp.mrc.ac.uk/ for details
            or contact: biohelp@hgmp.mrc.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .139
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="152h1"
                     /sex="male"
                     /tissue_type="blood"
                     /clone_lib="CGI-1"
                     /dev_stage="adult"
BASE COUNT       26 a      27 c      33 g      52 t       1 others
ORIGIN

  Query Match             100.0%;  Score 10;  DB 9;  Length 139;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db        122 GAGTTTTGTT 131
```

```
RESULT 36
HSPA33A6/c
LOCUS       HSPA33A6                139 bp    DNA     linear   STS 21-MAY-1998
DEFINITION  H.sapiens flow-sorted chromosome 6 HindIII fragment, SC6pA33A6,
            sequence tagged site.
ACCESSION   Z94241
VERSION     Z94241.1  GI:1945235
KEYWORDS    STS; single read.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 139)
  AUTHORS   Mungall,A.J., Huckle,E., Langford,C., Ross,M.T. and Hunt,S.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-APR-1997) The Sanger Centre, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail contact:
            humquery@sanger.ac.uk
COMMENT     Vector: pBSIISK+.
FEATURES             Location/Qualifiers
     source          1. .139
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="SC6pA33A6"
                     /sex="Female"
                     /tissue_type="EBV lymphoblastoid cell line"
                     /clone_lib="SC6pA"
                     /dev_stage="adult"
                     /note="The estimated purity of the flow-sorted chromosome
                     6 library is >97%"
BASE COUNT       52 a     31 c     22 g     34 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 11;  Length 139;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db        129 GAGTTTTGTT 120


RESULT 37
CLU55846
LOCUS       CLU55846                142 bp    DNA     linear   VRT 15-APR-2002
DEFINITION  Columba livia 16S ribosomal RNA gene, mitochondrial gene encoding
            mitochondrial rRNA, partial sequence.
ACCESSION   U55846
VERSION     U55846.1  GI:1305540
KEYWORDS    .
SOURCE      mitochondrion Columba livia (domestic pigeon)
  ORGANISM  Columba livia
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Columbiformes; Columbidae; Columba.
REFERENCE   1  (bases 1 to 142)
  AUTHORS   Parker,A. and Kornfield,I.
  TITLE     An improved amplification and sequencing strategy for phylogenetic
            studies using the mitochondrial large subunit rRNA gene
  JOURNAL   Genome 39 (4), 793-797 (1996)
  MEDLINE   96373168
    PUBMED  8776869
REFERENCE   2  (bases 1 to 142)
  AUTHORS   Parker,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-APR-1996) Alex Parker, University of Maine, Zoology,
            Orono, ME 04469-5751, USA
FEATURES             Location/Qualifiers
     source          1. .142
                     /organism="Columba livia"
                     /organelle="mitochondrion"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8932"
     rRNA            <1. .>142
                     /product="16S ribosomal RNA"
BASE COUNT       25 a     22 c     49 g     46 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 5;  Length 142;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db        115 GAGTTTTGTT 124


RESULT 38
BD053341
LOCUS       BD053341                142 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Sequence tag and encoded human protein.
ACCESSION   BD053341
VERSION     BD053341.1  GI:22656147
KEYWORDS    JP 2001269182-A/29587.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 142)
  AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
  TITLE     Sequence tag and encoded human protein
  JOURNAL   Patent: JP 2001269182-A 29587 02-OCT-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001269182-A/29587
            PD   02-OCT-2001
            PF   24-FEB-2000 JP 2000118773
            PR   26-FEB-1999 US    60/122487
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
            PI   JORDAN
            PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
            PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
              G06F15/40
            CC
            FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .142
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT       24 a     36 c     27 g     53 t      2 others
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 142;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         52 GAGTTTTGTT 61


RESULT 39
E50611
LOCUS       E50611                  144 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Betacellulin modification.
ACCESSION   E50611
VERSION     E50611.1  GI:18622103
KEYWORDS    JP 2000312591-A/11.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 144)
```

```
  AUTHORS   Ito,T., Kondo,M., Tanaka,Y., Kobayashi,M., Igarashi,K., Sasada,R.
            and Nishimura,H.
  TITLE     Betacellulin modification
  JOURNAL   Patent: JP 2000312591-A 11 14-NOV-2000;
            TAKEDA CHEM IND LTD
COMMENT     OS   Artificial Sequence
            PN   JP 2000312591-A/11
            PD   14-NOV-2000
            PF   08-DEC-1999 JP 1999348531
            PR
            PI   TAKASHI ITO,MITSUYO KONDO,YOKO TANAKA,MASAYUKI KOBAYASHI, PI
            KOICHI IGARASHI,
            PI   REIKO SASADA,HAJIME NISHIMURA
            PC   C12N15/09,A61K38/00,A61P3/10,C07K14/47,C12P21/02//(C12N15/09,
            PC   C12R1:91),
            PC   (C12P21/02,C12R1:19),C12N15/00,A61K37/02,(C12N15/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   source          1. .144
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .144
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT       34 a     33 c     43 g     34 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 144;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db        134 GAGTTTTGTT 143


RESULT 40
G26244/c
LOCUS       G26244                   144 bp    DNA     linear   STS 02-JUN-1996
DEFINITION  human STS TIGR-A004U39, sequence tagged site.
ACCESSION   G26244
VERSION     G26244.1  GI:1348476
KEYWORDS    STS; STS sequence; primer; sequence tagged site.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 144)
  AUTHORS   Hudson,T.
  TITLE     Whitehead Institute/MIT Center for Genome Research; Physically
            Mapped STSs
  JOURNAL   Unpublished (1995)
COMMENT
            Contact: Thomas Hudson
            Whitehead Institute/MIT Center for Genome Research
            Whitehead Institute for Biomedical Research
            9 Cambridge Center, Cambridge MA 02142 USA
            Tel: 617 252 1900
            Fax: 617 252 1902
            Email: thudson@genome.wi.mit.edu

            Primer A: GTCTACAAAAAGCAAAACAA
            Primer B: ACTCTAAGAGGTGTGCATATATA
            STS size: 144
            PCR Profile:
                  Presoak:
                  Denaturation:
                  Annealing: 56 degrees C
                  Polymerization:
                  PCR Cycles: 35
                  Thermal Cycler:
            Protocol:
                  Template: 10 ng
                  Primer: each 5 pM
                  dNTPs: each 4 nM
                  Taq Polymerase: 0.025 units/ul
                  Total Vol: 20 ul

               Buffer:
                  MgCl2: 1.5 mM
                  KCl: 50 mM
                  Tris-HCL: 10 mM
                  pH: 9.3

               Derived from dbEST (genbank accession Z38629).
FEATURES             Location/Qualifiers
     source          1. .144
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="28.1 cR from top of Chr18 linkage group"
     STS             1. .144
     primer_bind     1. .20
     primer_bind     complement(122. .144)
BASE COUNT       66 a     27 c     13 g     37 t      1 others
ORIGIN

  Query Match             100.0%;  Score 10;  DB 11;  Length 144;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         30 GAGTTTTGTT 21


RESULT 41
AG3UTRA41
LOCUS       AG3UTRA41                146 bp    mRNA    linear   INV 19-FEB-1998
DEFINITION  A.gambiae mRNA for 3'UTR of unknown protein, clone A41.
ACCESSION   Y08165
VERSION     Y08165.1  GI:1561557
KEYWORDS    unknown protein.
SOURCE      Anopheles gambiae (African malaria mosquito)
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
            Anopheles.
REFERENCE   1
  AUTHORS   Dimopoulos,G., Richman,A., della Torre,A., Kafatos,F.C. and
            Louis,C.
  TITLE     Identification and characterization of differentially expressed
            cDNAs of the vector mosquito, Anopheles gambiae
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 93 (23), 13066-13071 (1996)
  MEDLINE   97075119
   PUBMED   8917545
REFERENCE   2  (bases 1 to 146)
  AUTHORS   Richman,A.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-AUG-1996) A.M. Richman, European Molecular Biology
            Laboratory, DG Group, Meyerhofstrasse 1, 69117 Heidelberg, FRG
FEATURES             Location/Qualifiers
     source          1. .146
                     /organism="Anopheles gambiae"
                     /mol_type="mRNA"
                     /strain="G3"
                     /db_xref="taxon:7165"
                     /clone="A41"
                     /dev_stage="larval"
     3'UTR           <1. .>146
                     /note="unknown protein"
BASE COUNT       37 a     28 c     35 g     46 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 3;  Length 146;
```

```
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db        137 GAGTTTTGTT 146


RESULT 42
E50610
LOCUS       E50610                   147 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Betacellulin modification.
ACCESSION   E50610
VERSION     E50610.1  GI:18622102
KEYWORDS    JP 2000312591-A/10.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 147)
  AUTHORS   Ito,T., Kondo,M., Tanaka,Y., Kobayashi,M., Igarashi,K., Sasada,R.
            and Nishimura,H.
  TITLE     Betacellulin modification
  JOURNAL   Patent: JP 2000312591-A 10 14-NOV-2000;
            TAKEDA CHEM IND LTD
COMMENT     OS   Artificial Sequence
            PN   JP 2000312591-A/10
            PD   14-NOV-2000
            PF   08-DEC-1999 JP 1999348531
            PR
            PI   TAKASHI ITO,MITSUYO KONDO,YOKO TANAKA,MASAYUKI KOBAYASHI, PI
            KOICHI IGARASHI,
            PI   REIKO SASADA,HAJIME NISHIMURA
            PC   C12N15/09,A61K38/00,A61P3/10,C07K14/47,C12P21/02//(C12N15/09,
            PC   C12R1:91),
            PC   (C12P21/02,C12R1:19),C12N15/00,A61K37/02,(C12N15/00,C12R1:91)
            CC
            FH   Key             Location/Qualifiers
            FT   source          1. .147
            FT                   /organism='Artificial Sequence'.
FEATURES              Location/Qualifiers
     source           1. .147
                      /organism="synthetic construct"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:32630"
BASE COUNT       35 a     34 c     43 g     35 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 147;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db        134 GAGTTTTGTT 143


RESULT 43
AF462550/c
LOCUS       AF462550                 149 bp    DNA     linear   VRT 10-FEB-2002
DEFINITION  Hemibagrus nemurus clone PCTD3 microsatellite sequence.
ACCESSION   AF462550
VERSION     AF462550.1  GI:18643190
KEYWORDS    .
SOURCE      Hemibagrus nemurus (Asian redtail catfish)
  ORGANISM  Hemibagrus nemurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
            Bagridae; Hemibagrus.
REFERENCE   1  (bases 1 to 149)
  AUTHORS   Usmani,S. and Guan,T.S.
  TITLE     Identification and characterization of microsatellite markers in
            the Southeast Asian river catfish Mystus nemurus
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 149)
  AUTHORS   Usmani,S. and Guan,T.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-DEC-2001) Biology, Universiti Putra Malaysia, Faculty
            of Science and Environmental Studies, Serdang, Selangor 43400,
            Malaysia
FEATURES              Location/Qualifiers
     source           1. .149
                      /organism="Hemibagrus nemurus"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:156983"
                      /clone="PCTD3"
     repeat_region    8. .55
                      /note="microsatellite"
                      /rpt_type=tandem
     repeat_region    122. .141
                      /note="microsatellite"
                      /rpt_type=tandem
BASE COUNT       49 a     21 c     25 g     54 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 5;  Length 149;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         72 GAGTTTTGTT 63


RESULT 44
AX341152
LOCUS       AX341152                 149 bp    DNA     linear   PAT 10-JAN-2002
DEFINITION  Sequence 1399 from Patent WO0196388.
ACCESSION   AX341152
VERSION     AX341152.1  GI:18137134
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Jiang,Y., Harlocker,S.L. and Secrist,H.
  TITLE     Compositions and methods for the therapy and diagnosis of colon
            cancer
  JOURNAL   Patent: WO 0196388-A 1399 20-DEC-2001;
            CORIXA CORPORATION (US)
FEATURES              Location/Qualifiers
     source           1. .149
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
BASE COUNT       26 a     44 c     41 g     34 t      4 others
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 149;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db        100 GAGTTTTGTT 109


RESULT 45
BD056116
LOCUS       BD056116                 149 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Sequence tag and encoded human protein.
ACCESSION   BD056116
VERSION     BD056116.1  GI:22601722
KEYWORDS    JP 2001269182-A/32362.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

```
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 149)
  AUTHORS  Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
  TITLE    Sequence tag and encoded human protein
  JOURNAL  Patent: JP 2001269182-A 32362 02-OCT-2001;
           GENSET
COMMENT    OS   Homo sapiens (human)
           PN   JP 2001269182-A/32362
           PD   02-OCT-2001
           PF   24-FEB-2000 JP 2000118773
           PR   26-FEB-1999 US    60/122487
           PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
           PI   JORDAN
           PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
            C12N5/10,
           PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
             G06F15/40
           CC
           FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .149
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT       47 a     19 c     28 g     55 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 149;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         91 GAGTTTTGTT 100


RESULT 46
BD038248
LOCUS       BD038248                153 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Sequence tag and encoded human protein.
ACCESSION   BD038248
VERSION     BD038248.1  GI:22579990
KEYWORDS    JP 2001269182-A/14494.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 153)
  AUTHORS  Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
  TITLE    Sequence tag and encoded human protein
  JOURNAL  Patent: JP 2001269182-A 14494 02-OCT-2001;
           GENSET
COMMENT    OS   Homo sapiens (human)
           PN   JP 2001269182-A/14494
           PD   02-OCT-2001
           PF   24-FEB-2000 JP 2000118773
           PR   26-FEB-1999 US    60/122487
           PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
           PI   JORDAN
           PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
            C12N5/10,
           PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
             G06F15/40
           CC
           FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .153
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT       61 a     14 c     17 g     61 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 153;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db        125 GAGTTTTGTT 134


RESULT 47
AY033692
LOCUS       AY033692                154 bp    DNA     linear   PLN 18-JUN-2001
DEFINITION  Arabidopsis thaliana isolate RATHE1-58 SINE repeat sequence.
ACCESSION   AY033692
VERSION     AY033692.1  GI:14486212
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE  1  (bases 1 to 154)
  AUTHORS  Deragon,J.-M.
  TITLE    SINEs from Arabidopsis thaliana
  JOURNAL  Unpublished
REFERENCE  2  (bases 1 to 154)
  AUTHORS  Deragon,J.-M.
  TITLE    Direct Submission
  JOURNAL  Submitted (30-APR-2001) BIOMOVE, CNRS6547 University Blaise Pascal,
           24 Avenue des Landais, Aubiere 63177, France
FEATURES             Location/Qualifiers
     source          1. .154
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /isolate="RATHE1-58"
                     /db_xref="taxon:3702"
     repeat_region   1. .154
                     /rpt_family="SINE"
                     /rpt_type=dispersed
BASE COUNT       50 a     26 c     31 g     47 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 8;  Length 154;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db        120 GAGTTTTGTT 129


RESULT 48
G48354
LOCUS       G48354                  154 bp    DNA     linear   STS 26-MAR-1999
DEFINITION  SHGC-68482 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION   G48354
VERSION     G48354.1  GI:4529014
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 154)
  AUTHORS  Myers,R.M.
  TITLE    Human STSs (1999)
  JOURNAL  Unpublished (1999)
COMMENT
           Contact: Richard M. Myers
           Stanford Human Genome Center (SHGC)
           Stanford University School of Medicine
           Department of Genetics, M-344, Stanford, CA 94305, USA
           Tel: 4157259687
```

```
            Fax: 4157259689
            Email: myers@shgc.stanford.edu
            Primer A: TTTGCAACTTTTCTCAATCATTTT
            Primer B: TCATATCATTAAATGCAAATAGGCTT
            STS size: 129
            PCR Profile:
                    Initial incubation: 95 degrees C for 10 minutes
                    Denaturation:       94 degrees C for 30 seconds
                    Annealing:          60 degrees C for 30 seconds
                    Polymerization:     72 degrees C for 23 seconds
                    PCR Cycles:         30
                    Thermal Cycler:     Perkin Elmer 9700
            Protocol:
                    Template:                   25 ng
                    Primer:                     each 1 uM
                    dNTPs:                      each 200 uM
                    AmpliTaq Gold Polymerase:   0.07 units/ul
                    Total Vol:                  5 ul

            Buffer:
                    MgCl2:          2.5 mM
                    KCl:            50 mM
                    Tris-HCl:       10 mM
                    pH:             8.3

            BAC ends sequenced at TIGR from the CIT-HSP BAC library. Designed
            and developed at the Stanford Human Genome Center.
FEATURES                Location/Qualifiers
     source             1. .154
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /map="6"
                        /clone_lib="Human"
     STS                12. .140
     primer_bind        12. .35
     primer_bind        complement(115. .140)
BASE COUNT       46 a     19 c     18 g     71 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 11;  Length 154;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         71 GAGTTTTGTT 80


RESULT 49
AX072015/c
LOCUS       AX072015                155 bp    DNA     linear   PAT 25-JAN-2001
DEFINITION  Sequence 2487 from Patent WO0102568.
ACCESSION   AX072015
VERSION     AX072015.1  GI:12582366
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.,
            Kassam,A., Reinhard,C., Randazzo,F., Kennedy,G.C., Pot,D.,
            Lamson,G., Drmanac,R., Crkenjakov,R., Drmanac,S., Dickson,M.,
            Labat,I., Leshkowitiz,D., Kita,D., Garcia,V. and Strache-Crain,B.
  TITLE     Human genes and gene expression products
  JOURNAL   Patent: WO 0102568-A 2487 11-JAN-2001;
            CHIRON CORPORATION (US) ; HYSEQ, INC. (US)
FEATURES                Location/Qualifiers
     source             1. .155
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
BASE COUNT       26 a     59 c     58 g     12 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 155;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         56 GAGTTTTGTT 47


RESULT 50
AX111502/c
LOCUS       AX111502                155 bp    DNA     linear   PAT 29-MAY-2002
DEFINITION  Sequence 2235 from Patent WO0123604.
ACCESSION   AX111502
VERSION     AX111502.1  GI:13927794
KEYWORDS    .
SOURCE      Granulicatella adiacens
  ORGANISM  Granulicatella adiacens
            Bacteria; Firmicutes; Lactobacillales; Carnobacteriaceae;
            Granulicatella.
REFERENCE   1
  AUTHORS   Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
            Picard,F.J. and Roy,P.H.
  TITLE     Highly conserved genes and their use to generate probes and primers
            for detection of microorganisms
  JOURNAL   Patent: WO 0123604-A 2235 05-APR-2001;
            Infectio Diagnostic (I.D.I.) INC. (CA)
FEATURES                Location/Qualifiers
     source             1. .155
                        /organism="Granulicatella adiacens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:46124"
                        /note="ATCC 49175"
BASE COUNT       53 a     27 c     24 g     51 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 6;  Length 155;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db        132 GAGTTTTGTT 123


Search completed: January  2, 2004, 17:20:04
Job time : 2730 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 2, 2004, 16:15:00 ; Search time 285 Seconds
(without alignments)
94.717 Million cell updates/sec

Title: US-09-875-453B-5
Perfect score: 10
Sequence: 1 gagttttgtt 10

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	10	100.0	10	24	ABK29856	Cyclin D1 -30 to -
	2	10	100.0	10	24	ABK29946	Cyclin D1 promoter
c	3	10	100.0	12	23	ABH70569	Oligonucleotide pr
	4	10	100.0	12	23	ABH72581	Oligonucleotide pr
c	5	10	100.0	12	23	ABH77113	Oligonucleotide pr
	6	10	100.0	12	23	ABH89723	Oligonucleotide pr
	7	10	100.0	12	23	ABH93366	Oligonucleotide pr
c	8	10	100.0	12	23	ABH98987	Oligonucleotide pr
c	9	10	100.0	12	23	ABI01365	Oligonucleotide pr
	10	10	100.0	12	23	ABI04066	Oligonucleotide pr
c	11	10	100.0	12	23	ABI05890	Oligonucleotide pr
c	12	10	100.0	12	23	ABI06734	Oligonucleotide pr
c	13	10	100.0	12	23	ABI21217	Oligonucleotide pr
c	14	10	100.0	12	23	ABI26253	Oligonucleotide pr
c	15	10	100.0	12	23	ABI40629	Oligonucleotide pr
	16	10	100.0	12	23	ABI50929	Oligonucleotide pr
c	17	10	100.0	12	23	ABI52706	Oligonucleotide pr
c	18	10	100.0	12	23	ABI56602	Oligonucleotide pr
	19	10	100.0	12	23	ABI59137	Oligonucleotide pr
	20	10	100.0	12	23	ABI67458	Oligonucleotide pr
	21	10	100.0	13	23	ABC19684	Oligonucleotide SE
c	22	10	100.0	13	23	ABC19685	Oligonucleotide SE
	23	10	100.0	13	23	ABC24220	Oligonucleotide SE
c	24	10	100.0	13	23	ABC24221	Oligonucleotide SE
	25	10	100.0	13	23	ABC43378	Oligonucleotide SE
c	26	10	100.0	13	23	ABC43379	Oligonucleotide SE
	27	10	100.0	13	23	ABC55754	Oligonucleotide SE
c	28	10	100.0	13	23	ABC55755	Oligonucleotide SE
	29	10	100.0	13	23	ABF07590	Oligonucleotide SE
c	30	10	100.0	13	23	ABF07591	Oligonucleotide SE
	31	10	100.0	13	23	ABF32770	Oligonucleotide SE
c	32	10	100.0	13	23	ABF32771	Oligonucleotide SE
	33	10	100.0	13	23	ABF33996	Oligonucleotide SE
c	34	10	100.0	13	23	ABF33997	Oligonucleotide SE
	35	10	100.0	13	23	ABF52356	Oligonucleotide SE
c	36	10	100.0	13	23	ABF52357	Oligonucleotide SE
	37	10	100.0	13	23	ABF59822	Oligonucleotide SE
c	38	10	100.0	13	23	ABF59823	Oligonucleotide SE
	39	10	100.0	13	23	ABF75028	Oligonucleotide SE
c	40	10	100.0	13	23	ABF75029	Oligonucleotide SE
	41	10	100.0	13	23	ABF86708	Oligonucleotide SE
c	42	10	100.0	13	23	ABF86709	Oligonucleotide SE
	43	10	100.0	13	23	ABH46274	Oligonucleotide SE
c	44	10	100.0	13	23	ABH46275	Oligonucleotide SE
	45	10	100.0	13	23	ABH58528	Oligonucleotide SE
c	46	10	100.0	13	23	ABH58529	Oligonucleotide SE
	47	10	100.0	20	25	AAD49244	Human phospholipid
c	48	10	100.0	21	24	AAD47257	Human RELP gene-sp
c	49	10	100.0	21	25	ABZ21653	Human REG-like pro
c	50	10	100.0	23	21	AAA59606	PCR primer used to
	51	10	100.0	33	21	AAZ31480	Neisseria chimeric
c	52	10	100.0	39	20	AAX90439	Oligonucleotide SE
c	53	10	100.0	39	20	AAX55875	Human phospholipid
	54	10	100.0	42	21	AAA60144	Human APC gene var
c	55	10	100.0	47	21	AAZ66335	Human map-related
c	56	10	100.0	50	22	AAI78874	Human silent SNP c
	57	10	100.0	50	24	ABZ01038	Human leukocyte ge
	58	10	100.0	50	24	ABZ07400	Human leukocyte ge
c	59	10	100.0	51	22	AAI78875	Human silent SNP c
c	60	10	100.0	51	25	ABX54925	Bovine EST associa
c	61	10	100.0	60	24	ABN43708	Human spliced tran
c	62	10	100.0	60	24	ABN45888	Human spliced tran
	63	10	100.0	64	19	AAX11591	Human biallelic po
c	64	10	100.0	65	24	ABZ26555	Candida essential
c	65	10	100.0	65	24	ABZ26688	Candida essential
c	66	10	100.0	65	24	ABZ27498	Candida essential
	67	10	100.0	65	24	ABZ28643	Candida gene relat
	68	10	100.0	65	24	ABZ28646	Candida gene relat
	69	10	100.0	65	24	ABZ29704	Candida gene relat
	70	10	100.0	65	24	ABN53822	Mouse spliced tran
	71	10	100.0	76	25	ABX12141	Viral protein, CD4
	72	10	100.0	89	22	ABA18889	Human nervous syst
	73	10	100.0	105	22	AAK71245	Human immune/haema
c	74	10	100.0	107	24	ABS72358	Human gene trapped
	75	10	100.0	116	21	ABN81018	Shrimp polynucleot
	76	10	100.0	121	22	ABA81250	PSEN1 mutation cor
c	77	10	100.0	121	22	ABA81251	PSEN1 mutation cor
	78	10	100.0	125	22	AAI61553	Soybean 318013 reg
c	79	10	100.0	125	25	ABZ53498	Aspergillus oryzae
c	80	10	100.0	129	22	AAS42049	Genomic sequence #
c	81	10	100.0	129	22	AAS42050	Genomic sequence #

c 82	10	100.0	129	22	AAS42051	Genomic sequence #
c 83	10	100.0	133	21	AAC26256	Human secreted pro
c 84	10	100.0	135	25	ABX20595	Human GDP-mannose
85	10	100.0	136	24	ABV96877	Human pancreatic c
c 86	10	100.0	139	22	AAI61431	Soybean 240017 reg
c 87	10	100.0	141	25	ABX60747	Arabidopsis thalia
88	10	100.0	142	21	AAC29596	Human secreted pro
c 89	10	100.0	144	20	AAH85650	Human single nucle
90	10	100.0	144	21	AAA60807	Beta-cellulin mute
91	10	100.0	145	24	ABV97485	Human pancreatic c
92	10	100.0	147	21	AAA60806	Beta-cellulin mute
93	10	100.0	149	21	AAC32371	Human secreted pro
94	10	100.0	149	24	ABL37810	Human colon tumour
95	10	100.0	152	24	ABV99069	Human pancreatic c
c 96	10	100.0	153	21	AAC54565	Arabidopsis thalia
97	10	100.0	153	21	AAC14503	Human secreted pro
c 98	10	100.0	155	21	AAA41630	Human secreted exp
c 99	10	100.0	155	22	AAH02242	Abiotrophia adiace
c 100	10	100.0	155	22	AAF66731	Novel human polynu
c 101	10	100.0	155	25	ABX26803	Human GDP-mannose
c 102	10	100.0	160	24	ABK93226	Human prostate spe
103	10	100.0	162	17	AAT12398	3'-Flanking sequen
104	10	100.0	163	20	AAH86394	Human single nucle
c 105	10	100.0	163	23	AAS48337	Enterococcus faeca
c 106	10	100.0	166	21	AAA41702	Human secreted exp
107	10	100.0	168	21	AAC45463	Arabidopsis thalia
108	10	100.0	168	24	ABL80513	Human ovarian canc
109	10	100.0	172	22	ABA71227	Human foetal liver
110	10	100.0	172	22	ABA37536	Probe #16002 for g
111	10	100.0	172	22	AAK19527	Human brain expres
112	10	100.0	172	22	AAK45522	Human bone marrow
113	10	100.0	172	22	AAI25304	Probe #15237 for g
114	10	100.0	172	22	AAI51463	Probe #20149 used
115	10	100.0	172	23	ABS45209	Human liver single
116	10	100.0	172	24	ABS19791	Human genome-deriv
c 117	10	100.0	173	21	AAC15702	Human secreted pro
118	10	100.0	174	25	ABX27418	Human GDP-mannose
c 119	10	100.0	175	21	AAC16518	Human secreted pro
120	10	100.0	177	20	AAX99540	Nucleic acid seque
c 121	10	100.0	179	16	AAT24213	Human gene signatu
122	10	100.0	179	21	AAC15815	Human secreted pro
123	10	100.0	183	24	ABN77893	Human ORF2840 cDNA
c 124	10	100.0	189	22	AAK79197	Human immune/haema
125	10	100.0	192	25	ABZ73238	Rice leaf EST, SEQ
126	10	100.0	193	24	ABL93218	Rat metastatic tum
c 127	10	100.0	198	23	ABV08566	Human prostate exp
128	10	100.0	199	24	ABQ55138	Human ovarian anti
c 129	10	100.0	202	15	AAQ76822	Human genome fragm
c 130	10	100.0	203	22	AAK61028	Human immune/haema
c 131	10	100.0	204	24	ABN65545	Human cancer relat
c 132	10	100.0	205	25	ABX61258	Arabidopsis thalia
133	10	100.0	207	24	ABN20723	Human ORFX polynuc
134	10	100.0	209	22	ABA19431	Human nervous syst
135	10	100.0	209	22	ABA19433	Human nervous syst
c 136	10	100.0	214	20	AAV87460	EST clone BP870.
137	10	100.0	214	21	AAC09616	Human secreted pro
c 138	10	100.0	214	21	AAC26401	Human secreted pro
139	10	100.0	219	22	ABA74130	Human foetal liver
140	10	100.0	219	22	ABA39145	Probe #17611 for g
141	10	100.0	219	22	AAK22583	Human brain expres
142	10	100.0	219	22	AAK48752	Human bone marrow
143	10	100.0	219	22	AAI54582	Probe #23268 used
144	10	100.0	219	23	ABS48426	Human liver single
145	10	100.0	219	24	ABS22465	Human genome-deriv
c 146	10	100.0	221	22	AAL24745	Human breast cance
147	10	100.0	222	20	AAX37071	Human cdc37 nuclei
148	10	100.0	222	20	AAX36989	Human cdc37 nuclei
c 149	10	100.0	226	22	AAK58860	Human immune/haema
150	10	100.0	226	24	ABL51117	Human DL intron 11

ALIGNMENTS

RESULT 1
ABK29856
ID ABK29856 standard; DNA; 10 BP.
XX
AC ABK29856;
XX
DT 23-APR-2002 (first entry)
XX
DE Cyclin D1 -30 to -21 wild type sequence.
XX
KW Cyclin D1 promoter; CD40L promoter; hepatitis B virus promoter;
KW HBV promoter; vancomycin-resistant enterococci promoter; VRE promoter;
KW vanH promoter; androgen receptor promoter; AR promoter;
KW human epidermal growth factor receptor 2 promoter; her2 promoter;
KW beta lactamase promoter; Bla promoter; transgene; cancer; breast cancer;
KW colon cancer; immunological disorder; prostate cancer; cytostatic;
KW autoimmune disease; HBV pre-S promoter; HBV-X promoter;
KW Enterococcus infection; immunosuppressive; antibacterial; antiviral;
KW gene expression modulator; multiple sclerosis; MS;
KW chronic hepatic insufficiency; cirrhosis; hepatocellular carcinoma;
KW systematic lupus erythematosus; SLE; graft-vs-host disease; GVHD;
KW familial adenomatous polyposis; rheumatoid arthritis; PCR; primer;
KW transgenic; ss.
XX
OS Homo sapiens.
XX
PN WO200194600-A2.
XX
PD 13-DEC-2001.
XX
PF 06-JUN-2001; 2001WO-US18343.
XX
PR 06-JUN-2000; 2000US-209549P.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Kim JP, Starr DB, Tam AW, Laurance ME, Michelotti EF, Velligan MD;
PI Latour DR, Thomas RL, Kongpachith A, Sheppard LT, Lim MY;
PI Bruice TW;
XX
DR WPI; 2002-130595/17.
XX
PT New nucleic acid regulatory sequences, which are able to regulate
PT expression of a gene operably linked to a promoter, useful for
PT regulating the expression of transgenes and for treating e.g., cancer
PT and immunological diseases -
XX
PS Claim 2; Page 58; 95pp; English.
XX
CC The invention describes an isolated nucleic acid regulatory sequence for
CC a cyclin D1 promoter, a CD40L promoter, vancomycin-resistant enterococci
CC (VRE) promoter, an HBV promoter, androgen receptor (AR) promoter, Human
CC epidermal growth factor receptor 2 (HER2) promoter, or a beta lactamase
CC (Bla) promoter. Transcription regulatory sequences may be used to
CC regulate expression of the endogenous, autologous or heterologous genes
CC operably linked to the promoter, and may be incorporated into
CC heterologous nucleic acid constructs for use in regulated expression of
CC transgenes. Regulated expression of cyclin D1 can be used in cancer
CC therapies, such as breast, colon or pancreatic cancers and familial
CC adenomatous polyposis. Regulation of the activity of CD40L gene promoter
CC may be used in the treatment of immunological disorders, such as
CC autoimmune diseases e.g. multiple sclerosis (MS), systematic lupus
CC erythematosus (SLE), graft-vs-host disease (GVHD) and rheumatoid
CC arthritis. Regulated expression of genes under the control of the HBV
CC (hepatitis B)-specific core, pre-S and X promoters can be used in the
CC therapy of HBV disease, chronic hepatic insufficiency, cirrhosis,
CC hepatocellular carcinoma, and in the regulated expression of liver
CC cell-specific genes. Regulated expression of the vanH gene promoter can
CC be used in treatment of Enterococcus infection, while regulated
CC expression of the androgen receptor gene can be used in the treatment of
CC prostate cancer. This sequence represents a primer used in the invention
CC to determine the functions of regions within the selected promoters,

```
CC   described in the method of the invention.
XX
SQ   Sequence 10 BP; 1 A; 0 C; 3 G; 6 T; 0 other;

  Query Match           100.0%;  Score 10;  DB 24;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          1 GAGTTTTGTT 10


RESULT 2
ABK29946
ID   ABK29946 standard; DNA; 10 BP.
XX
AC   ABK29946;
XX
DT   23-APR-2002  (first entry)
XX
DE   Cyclin D1 promoter -30 to -21 region, mutant #1.
XX
KW   Cyclin D1 promoter; CD40L promoter; hepatitis B virus promoter;
KW   HBV promoter; vancomycin-resistant enterococci promoter; VRE promoter;
KW   vanH promoter; androgen receptor promoter; AR promoter;
KW   human epidermal growth factor receptor 2 promoter; her2 promoter;
KW   beta lactamase promoter; Bla promoter; transgene; cancer; breast cancer;
KW   colon cancer; immunological disorder; prostate cancer; cytostatic;
KW   autoimmune disease; HBV pre-S promoter; HBV-X promoter;
KW   Enterococcus infection; immunosuppressive; antibacterial; antiviral;
KW   gene expression modulator; multiple sclerosis; MS;
KW   chronic hepatic insufficiency; cirrhosis; hepatocellular carcinoma;
KW   systematic lupus erythematosus; SLE; graft-vs-host disease; GVHD;
KW   familial adenomatous polyposis; rheumatoid arthritis; PCR; primer;
KW   mutant; transgenic; ds.
XX
OS   Homo sapiens.
XX
PN   WO200194600-A2.
XX
PD   13-DEC-2001.
XX
PF   06-JUN-2001; 2001WO-US18343.
XX
PR   06-JUN-2000; 2000US-209549P.
XX
PA   (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI   Kim JP,  Starr DB,  Tam AW,  Laurance ME,  Michelotti EF,  Velligan MD;
PI   Latour DR,  Thomas RL,  Kongpachith A,  Sheppard LT,  Lim MY;
PI   Bruice TW;
XX
DR   WPI; 2002-130595/17.
XX
PT   New nucleic acid regulatory sequences, which are able to regulate
PT   expression of a gene operably linked to a promoter, useful for
PT   regulating the expression of transgenes and for treating e.g., cancer
PT   and immunological diseases  -
XX
PS   Example 1; Page 36; 95pp; English.
XX
CC   The invention describes an isolated nucleic acid regulatory sequence for
CC   a cyclin D1 promoter, a CD40L promoter, vancomycin-resistant enterococci
CC   (VRE) promoter, an HBV promoter, androgen receptor (AR) promoter, Human
CC   epidermal growth factor receptor 2 (HER2) promoter, or a beta lactamase
CC   (Bla) promoter. Transcription regulatory sequences may be used to
CC   regulate expression of the endogenous, autologous or heterologous genes
CC   operably linked to the promoter, and may be incorporated into
CC   heterologous nucleic acid constructs for use in regulated expression of
CC   transgenes. Regulated expression of cyclin D1 can be used in cancer
CC   therapies, such as breast, colon or pancreatic cancers and familial
CC   adenomatous polyposis. Regulation of the activity of CD40L gene promoter
CC   may be used in the treatment of immunological disorders, such as
CC   autoimmune diseases e.g. multiple sclerosis (MS), systematic lupus
CC   erythematosus (SLE), graft-vs-host disease (GVHD) and rheumatoid
CC   arthritis. Regulated expression of genes under the control of the HBV
CC   (hepatitis B)-specific core, pre-S and X promoters can be used in the
CC   therapy of HBV disease, chronic hepatic insufficiency, cirrhosis,
CC   hepatocellular carcinoma, and in the regulated expression of liver
CC   cell-specific genes. Regulated expression of the vanH gene promoter can
CC   be used in treatment of Enterococcus infection, while regulated
CC   expression of the androgen receptor gene can be used in the treatment of
CC   prostate cancer. This sequence represents a mutated promoter region used
CC   in the invention to determine the regulatory regions involved in gene
CC   expression, described in the method of the invention.
XX
SQ   Sequence 10 BP; 1 A; 0 C; 3 G; 6 T; 0 other;

  Query Match           100.0%;  Score 10;  DB 24;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          1 GAGTTTTGTT 10


RESULT 3
ABH70569/c
ID   ABH70569 standard; DNA; 12 BP.
XX
AC   ABH70569;
XX
DT   22-FEB-2002  (first entry)
XX
DE   Oligonucleotide primer SEQ ID NO 270546 for detecting SNP TSC0002178.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 270546; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
```

```
XX
SQ   Sequence 12 BP; 6 A; 5 C; 0 G; 1 T; 0 other;

  Query Match            100.0%;  Score 10;  DB 23;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         10 GAGTTTTGTT 1


RESULT 4
ABH72581
ID   ABH72581 standard; DNA; 12 BP.
XX
AC   ABH72581;
XX
DT   22-FEB-2002  (first entry)
XX
DE   Oligonucleotide primer SEQ ID NO 272566 for detecting SNP TSC0002861.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 272566; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 12 BP; 2 A; 0 C; 3 G; 7 T; 0 other;

  Query Match            100.0%;  Score 10;  DB 23;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          2 GAGTTTTGTT 11


RESULT 5
```

```
ABH77113/c
ID   ABH77113 standard; DNA; 12 BP.
XX
AC   ABH77113;
XX
DT   22-FEB-2002  (first entry)
XX
DE   Oligonucleotide primer SEQ ID NO 277106 for detecting SNP TSC0004385.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 277106; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 12 BP; 7 A; 3 C; 0 G; 2 T; 0 other;

  Query Match            100.0%;  Score 10;  DB 23;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         11 GAGTTTTGTT 2


RESULT 6
ABH89723
ID   ABH89723 standard; DNA; 12 BP.
XX
AC   ABH89723;
XX
DT   22-FEB-2002  (first entry)
XX
DE   Oligonucleotide primer SEQ ID NO 289716 for detecting SNP TSC0014062.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
```

OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 289716; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 12 BP; 2 A; 0 C; 3 G; 7 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 1 GAGTTTTGTT 10

RESULT 7
ABH93366
ID ABH93366 standard; DNA; 12 BP.
XX
AC ABH93366;
XX
DT 22-FEB-2002 (first entry)
XX
DE Oligonucleotide primer SEQ ID NO 293359 for detecting SNP TSC0015581.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 293359; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 12 BP; 1 A; 0 C; 3 G; 8 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 2 GAGTTTTGTT 11

RESULT 8
ABH98987/c
ID ABH98987 standard; DNA; 12 BP.
XX
AC ABH98987;
XX
DT 22-FEB-2002 (first entry)
XX
DE Oligonucleotide primer SEQ ID NO 298980 for detecting SNP TSC0018381.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 298980; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a

```
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 12 BP; 8 A; 3 C; 0 G; 1 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         11 GAGTTTTGTT 2


RESULT 9
ABI01365/c
ID   ABI01365 standard; DNA; 12 BP.
XX
AC   ABI01365;
XX
DT   22-FEB-2002. (first entry)
XX
DE   Oligonucleotide primer SEQ ID NO 301338 for detecting SNP TSC0019457.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 301338; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 12 BP; 8 A; 3 C; 0 G; 1 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         10 GAGTTTTGTT 1


RESULT 10
ABI04066
ID   ABI04066 standard; DNA; 12 BP.
XX
AC   ABI04066;
XX
DT   22-FEB-2002  (first entry)
XX
DE   Oligonucleotide primer SEQ ID NO 304039 for detecting SNP TSC0020755.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 304039; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 12 BP; 2 A; 0 C; 3 G; 7 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          1 GAGTTTTGTT 10


RESULT 11
ABI05890/c
ID   ABI05890 standard; DNA; 12 BP.
XX
AC   ABI05890;
XX
```

DT 22-FEB-2002 (first entry)
XX
DE Oligonucleotide primer SEQ ID NO 305863 for detecting SNP TSC0021676.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 305863; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 12 BP; 8 A; 3 C; 0 G; 1 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 12 GAGTTTTGTT 3

RESULT 12
ABI06734/c
ID ABI06734 standard; DNA; 12 BP.
XX
AC ABI06734;
XX
DT 22-FEB-2002 (first entry)
XX
DE Oligonucleotide primer SEQ ID NO 306707 for detecting SNP TSC0022140.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 306707; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 12 BP; 7 A; 3 C; 0 G; 2 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 12 GAGTTTTGTT 3

RESULT 13
ABI21217/c
ID ABI21217 standard; DNA; 12 BP.
XX
AC ABI21217;
XX
DT 22-FEB-2002 (first entry)
XX
DE Oligonucleotide primer SEQ ID NO 321190 for detecting SNP TSC0030097.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine

PT methylation status -
XX
PS Claim 1; SEQ ID 321190; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 12 BP; 7 A; 4 C; 0 G; 1 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 12 GAGTTTTGTT 3

RESULT 14
ABI26253/c
ID ABI26253 standard; DNA; 12 BP.
XX
AC ABI26253;
XX
DT 22-FEB-2002 (first entry)
XX
DE Oligonucleotide primer SEQ ID NO 326226 for detecting SNP TSC0032965.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 326226; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 12 BP; 7 A; 4 C; 0 G; 1 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 12 GAGTTTTGTT 3

RESULT 15
ABI40629/c
ID ABI40629 standard; DNA; 12 BP.
XX
AC ABI40629;
XX
DT 22-FEB-2002 (first entry)
XX
DE Oligonucleotide primer SEQ ID NO 340602 for detecting SNP TSC0041606.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 340602; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 12 BP; 6 A; 3 C; 0 G; 3 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 12 GAGTTTTGTT 3

```
RESULT 16
ABI50929
ID   ABI50929 standard; DNA; 12 BP.
XX
AC   ABI50929;
XX
DT   22-FEB-2002  (first entry)
XX
DE   Oligonucleotide primer SEQ ID NO 350902 for detecting SNP TSC0046964.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 350902; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 12 BP; 3 A; 0 C; 3 G; 6 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          1 GAGTTTTGTT 10


RESULT 17
ABI52706/c
ID   ABI52706 standard; DNA; 12 BP.
XX
AC   ABI52706;
XX
DT   22-FEB-2002  (first entry)
XX
DE   Oligonucleotide primer SEQ ID NO 352679 for detecting SNP TSC0048031.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 352679; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 12 BP; 6 A; 4 C; 0 G; 2 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         10 GAGTTTTGTT 1


RESULT 18
ABI56602/c
ID   ABI56602 standard; DNA; 12 BP.
XX
AC   ABI56602;
XX
DT   22-FEB-2002  (first entry)
XX
DE   Oligonucleotide primer SEQ ID NO 356575 for detecting SNP TSC0050201.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
```

PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 356575; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 12 BP; 7 A; 3 C; 0 G; 2 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTGTT 10
||||||||||
Db 12 GAGTTTGTT 3

RESULT 19
ABI59137
ID ABI59137 standard; DNA; 12 BP.
XX
AC ABI59137;
XX
DT 22-FEB-2002 (first entry)
XX
DE Oligonucleotide primer SEQ ID NO 359110 for detecting SNP TSC0008874.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 359110; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 12 BP; 2 A; 0 C; 3 G; 7 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTGTT 10
||||||||||
Db 3 GAGTTTGTT 12

RESULT 20
ABI67458
ID ABI67458 standard; DNA; 12 BP.
XX
AC ABI67458;
XX
DT 22-FEB-2002 (first entry)
XX
DE Oligonucleotide primer SEQ ID NO 367431 for detecting SNP TSC0006664.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 367431; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 12 BP; 1 A; 0 C; 3 G; 8 T; 0 other;

```
  Query Match             100.0%;  Score 10;  DB 23;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          1 GAGTTTTGTT 10


RESULT 21
ABC19684
ID   ABC19684 standard; DNA; 13 BP.
XX
AC   ABC19684;
XX
DT   20-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 19701 for detecting SNP TSC0004079.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 19701; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13 BP; 1 A; 0 C; 4 G; 8 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          1 GAGTTTTGTT 10


RESULT 22
ABC19685/c
ID   ABC19685 standard; DNA; 13 BP.
XX
AC   ABC19685;
XX
DT   20-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 19702 for detecting SNP TSC0004079.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 19702; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13 BP; 8 A; 4 C; 0 G; 1 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         13 GAGTTTTGTT 4


RESULT 23
ABC24220
ID   ABC24220 standard; DNA; 13 BP.
XX
AC   ABC24220;
XX
DT   20-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 24237 for detecting SNP TSC0005731.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
```

PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 24237; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 13 BP; 3 A; 0 C; 4 G; 6 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 1 GAGTTTTGTT 10

RESULT 24
ABC24221/c
ID ABC24221 standard; DNA; 13 BP.
XX
AC ABC24221;
XX
DT 20-FEB-2002 (first entry)
XX
DE Oligonucleotide SEQ ID NO 24238 for detecting SNP TSC0005731.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 24238; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 13 BP; 6 A; 4 C; 0 G; 3 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 13 GAGTTTTGTT 4

RESULT 25
ABC43378
ID ABC43378 standard; DNA; 13 BP.
XX
AC ABC43378;
XX
DT 21-FEB-2002 (first entry)
XX
DE Oligonucleotide SEQ ID NO 43395 for detecting SNP TSC0012841.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 43395; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The

```
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13 BP; 3 A; 0 C; 4 G; 6 T; 0 other;

  Query Match            100.0%;  Score 10;  DB 23;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          1 GAGTTTTGTT 10


RESULT 26
ABC43379/c
ID   ABC43379 standard; DNA; 13 BP.
XX
AC   ABC43379;
XX
DT   21-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 43396 for detecting SNP TSC0012841.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 43396; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13 BP; 6 A; 4 C; 0 G; 3 T; 0 other;

  Query Match            100.0%;  Score 10;  DB 23;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         13 GAGTTTTGTT 4


RESULT 27
ABC55754
ID   ABC55754 standard; DNA; 13 BP.
XX
AC   ABC55754;
XX
DT   21-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 55771 for detecting SNP TSC0015193.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 55771; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13 BP; 2 A; 0 C; 4 G; 7 T; 0 other;

  Query Match            100.0%;  Score 10;  DB 23;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          4 GAGTTTTGTT 13


RESULT 28
ABC55755/c
ID   ABC55755 standard; DNA; 13 BP.
XX
AC   ABC55755;
XX
DT   21-FEB-2002  (first entry)
XX
```

DE Oligonucleotide SEQ ID NO 55772 for detecting SNP TSC0015193.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 55772; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 13 BP; 7 A; 4 C; 0 G; 2 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 10 GAGTTTTGTT 1

RESULT 29
ABF07590
ID ABF07590 standard; DNA; 13 BP.
XX
AC ABF07590;
XX
DT 21-FEB-2002 (first entry)
XX
DE Oligonucleotide SEQ ID NO 107587 for detecting SNP TSC0026938.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 107587; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 13 BP; 2 A; 0 C; 5 G; 6 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 4 GAGTTTTGTT 13

RESULT 30
ABF07591/c
ID ABF07591 standard; DNA; 13 BP.
XX
AC ABF07591;
XX
DT 21-FEB-2002 (first entry)
XX
DE Oligonucleotide SEQ ID NO 107588 for detecting SNP TSC0026938.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX

```
PS   Claim 1; SEQ ID 107588; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13 BP; 6 A; 5 C; 0 G; 2 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         10 GAGTTTTGTT 1


RESULT 31
ABF32770
ID   ABF32770 standard; DNA; 13 BP.
XX
AC   ABF32770;
XX
DT   21-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 132767 for detecting SNP TSC0033108.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 132767; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13 BP; 1 A; 0 C; 3 G; 8 T; 1 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          1 GAGTTTTGTT 10


RESULT 32
ABF32771/c
ID   ABF32771 standard; DNA; 13 BP.
XX
AC   ABF32771;
XX
DT   21-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 132768 for detecting SNP TSC0033108.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 132768; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13 BP; 8 A; 3 C; 0 G; 1 T; 1 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         13 GAGTTTTGTT 4
```

```
RESULT 33
ABF33996
ID   ABF33996 standard; DNA; 13 BP.
XX
AC   ABF33996;
XX
DT   21-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 133993 for detecting SNP TSC0033412.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 133993; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13 BP; 3 A; 0 C; 3 G; 7 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          1 GAGTTTTGTT 10


RESULT 34
ABF33997/c
ID   ABF33997 standard; DNA; 13 BP.
XX
AC   ABF33997;
XX
DT   21-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 133994 for detecting SNP TSC0033412.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 133994; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13 BP; 7 A; 3 C; 0 G; 3 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         13 GAGTTTTGTT 4


RESULT 35
ABF52356
ID   ABF52356 standard; DNA; 13 BP.
XX
AC   ABF52356;
XX
DT   21-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 152353 for detecting SNP TSC0038490.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
```

PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 152353; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 13 BP; 1 A; 0 C; 4 G; 8 T; 0 other;

Query Match . 100.0%; Score 10; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 1 GAGTTTTGTT 10

RESULT 36
ABF52357/c
ID ABF52357 standard; DNA; 13 BP.
XX
AC ABF52357;
XX
DT 21-FEB-2002 (first entry)
XX
DE Oligonucleotide SEQ ID NO 152354 for detecting SNP TSC0038490.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 152354; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 13 BP; 8 A; 4 C; 0 G; 1 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 13 GAGTTTTGTT 4

RESULT 37
ABF59822
ID ABF59822 standard; DNA; 13 BP.
XX
AC ABF59822;
XX
DT 21-FEB-2002 (first entry)
XX
DE Oligonucleotide SEQ ID NO 159819 for detecting SNP TSC0040226.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 159819; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 13 BP; 3 A; 0 C; 3 G; 7 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 13;

```
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          4 GAGTTTTGTT 13


RESULT 38
ABF59823/c
ID   ABF59823 standard; DNA; 13 BP.
XX
AC   ABF59823;
XX
DT   21-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 159820 for detecting SNP TSC0040226.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 159820; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13 BP; 7 A; 3 C; 0 G; 3 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         10 GAGTTTTGTT 1


RESULT 39
ABF75028
ID   ABF75028 standard; DNA; 13 BP.
XX
AC   ABF75028;
XX
DT   22-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 175025 for detecting SNP TSC0043506.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 175025; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13 BP; 2 A; 0 C; 4 G; 7 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          1 GAGTTTTGTT 10


RESULT 40
ABF75029/c
ID   ABF75029 standard; DNA; 13 BP.
XX
AC   ABF75029;
XX
DT   22-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 175026 for detecting SNP TSC0043506.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
```

PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 175026; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 13 BP; 7 A; 4 C; 0 G; 2 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 13 GAGTTTTGTT 4

RESULT 41
ABF86708
ID ABF86708 standard; DNA; 13 BP.
XX
AC ABF86708;
XX
DT 22-FEB-2002 (first entry)
XX
DE Oligonucleotide SEQ ID NO 186705 for detecting SNP TSC0046017.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 186705; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 13 BP; 1 A; 0 C; 4 G; 8 T; 0 other;

Query Match 100.0%; Score 10; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 2 GAGTTTTGTT 11

RESULT 42
ABF86709/c
ID ABF86709 standard; DNA; 13 BP.
XX
AC ABF86709;
XX
DT 22-FEB-2002 (first entry)
XX
DE Oligonucleotide SEQ ID NO 186706 for detecting SNP TSC0046017.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 186706; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and

```
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13 BP; 8 A; 4 C; 0 G; 1 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         12 GAGTTTTGTT 3


RESULT 43
ABH46274
ID   ABH46274 standard; DNA; 13 BP.
XX
AC   ABH46274;
XX
DT   22-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 246251 for detecting SNP TSC0060177.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 246251; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13 BP; 2 A; 0 C; 5 G; 6 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          1 GAGTTTTGTT 10


RESULT 44
ABH46275/c
ID   ABH46275 standard; DNA; 13 BP.
XX
AC   ABH46275;
XX
DT   22-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 246252 for detecting SNP TSC0060177.
XX
KW   SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW   peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW   central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS   Homo sapiens.
XX
PN   WO200177384-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-IB00713.
XX
PR   07-APR-2000; 2000DE-1019173.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2001-657177/75.
XX
PT   Set of oligonucleotides, useful for diagnosis and cell typing, is
PT   designed to detect single nucleotide polymorphisms and cytosine
PT   methylation status  -
XX
PS   Claim 1; SEQ ID 246252; 29pp + Sequence Listing; German.
XX
CC   This invention describes novel oligonucleotide primers or peptide nucleic
CC   acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC   and cytosine methylation status in chemically pretreated genomic DNA. The
CC   oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC   range of diseases including immune system, gastrointestinal, respiratory,
CC   central nervous system, cardiovascular and metabolic disorders. The
CC   oligomers are also used for detecting cell type differentiation.
CC   ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC   ABI00010-ABI82073 represent the oligomers described in the invention.
CC   NOTE: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 13 BP; 6 A; 5 C; 0 G; 2 T; 0 other;

  Query Match             100.0%;  Score 10;  DB 23;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         13 GAGTTTTGTT 4


RESULT 45
ABH58528
ID   ABH58528 standard; DNA; 13 BP.
XX
AC   ABH58528;
XX
DT   22-FEB-2002  (first entry)
XX
DE   Oligonucleotide SEQ ID NO 258505 for detecting SNP TSC0062854.
XX
```

KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 258505; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 13 BP; 2 A; 0 C; 4 G; 6 T; 1 other;

Query Match 100.0%; Score 10; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 1 GAGTTTTGTT 10

RESULT 46
ABH58529/c
ID ABH58529 standard; DNA; 13 BP.
XX
AC ABH58529;
XX
DT 22-FEB-2002 (first entry)
XX
DE Oligonucleotide SEQ ID NO 258506 for detecting SNP TSC0062854.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-IB00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 258506; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 13 BP; 6 A; 4 C; 0 G; 2 T; 1 other;

Query Match 100.0%; Score 10; DB 23; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 13 GAGTTTTGTT 4

RESULT 47
AAD49244
ID AAD49244 standard; DNA; 20 BP.
XX
AC AAD49244;
XX
DT 07-MAR-2003 (first entry)
XX
DE Human phospholipid scramblase I antisense oligo, ISIS #120455.
XX
KW Human; antisense; phospholipid scramblase I; immune disorder; cancer;
KW inflammation; hyperproliferative; antisense therapy; phosphorothioate;
KW ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /*tag= a
FT /mod_base= OTHER
FT /note= "Phosphorothioate backbone"
FT modified_base 1..5
FT /*tag= b
FT /mod_base= OTHER
FT /note= "2'methoxyethyl nucleotides"
FT modified_base 16..20
FT /*tag= c
FT /mod_base= OTHER
FT /note= "2'methoxyethyl nucleotides"
FT modified_base 17..18
FT /*tag= d
FT /mod_base= m5c
XX
PN WO200281495-A1.
XX

PD 17-OCT-2002.
XX
PF 02-APR-2002; 2002WO-US10529.
XX
PR 05-APR-2001; 2001US-0828344.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Bennett CF, Wyatt JR;
XX
DR WPI; 2003-058495/05.
XX
PT Novel antisense compounds targeted to nucleic acids encoding
PT phospholipid scramblase I, for modulating gene expression and treating
PT inflammation, immune disorders and hyperproliferative conditions e.g.
PT cancer -
XX
PS Example 15; Page 76; 131pp; English.
XX
CC The invention relates to an antisense compound targetted to a nucleic
CC acid molecule encoding phospholipid scramblase I and which specifically
CC hybridises with and inhibits the expression of phospholipid scramblase I,
CC or which hybridises with at least an 8-nucleobase portion of an active
CC site on a nucleic acid molecule encoding phospholipid scramblase I. The
CC invention is useful for inhibiting the expression of human phospholipid
CC scramblase I in cells or tissues and for treating an animal having a
CC disease or condition associated with phospholipid scramblase I, such as
CC inflammation, an immune disorder and a hyperproliferative condition, e.g.
CC cancer. The invention is useful for diagnostics, therapeutics and as
CC research reagent. The present sequence is human phospholipid scramblase I
CC antisense oligonucleotide.
XX
SQ Sequence 20 BP; 2 A; 2 C; 5 G; 11 T; 0 other;

Query Match 100.0%; Score 10; DB 25; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 4 GAGTTTTGTT 13

RESULT 48
AAD47257/c
ID AAD47257 standard; DNA; 21 BP.
XX
AC AAD47257;
XX
DT 24-FEB-2003 (first entry)
XX
DE Human RELP gene-specific PCR primer #17.
XX
KW Human; REG-like protein; RELP; tumour; cancer; therapy; PCR; primer; ss.
XX
OS Homo sapiens.
XX
PN EP1241269-A2.
XX
PD 18-SEP-2002.
XX
PF 15-MAR-2002; 2002EP-0251876.
XX
PR 16-MAR-2001; 2001US-276414P.
XX
PA (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
XX
PI Heiskala M;
XX
DR WPI; 2002-684095/74.
XX
PT Detecting the presence of a tumor comprises detecting the concentration
PT of a Reg Like Protein or the presence or quantity of a nucleic acid
PT encoding it -
XX
PS Claim 7; Page 9; 26pp; English.
XX
CC The invention relates to a method for detecting REG-like protein (RELP)
CC and its nucleic acid sequence. The method is useful for detecting the
CC presence of a tumour. Kits comprising an antibody specific for RELP and
CC reagents for detecting the antibody, or a nucleic acid complementary to
CC a portion of a nucleic acid encoding RELP, are useful for identifying
CC the presence of cancer, characterise the cancer, or monitor the course
CC of treatment of cancer. The present sequence is a PCR primer used for
CC amplifying human RELP gene. This sequence is used to illustrate the
CC method of the invention.
XX
SQ Sequence 21 BP; 9 A; 7 C; 1 G; 4 T; 0 other;

Query Match 100.0%; Score 10; DB 24; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 21 GAGTTTTGTT 12

RESULT 49
ABZ21653/c
ID ABZ21653 standard; DNA; 21 BP.
XX
AC ABZ21653;
XX
DT 26-FEB-2003 (first entry)
XX
DE Human REG-like protein (RELP) specific PCR primer SEQ ID NO:28.
XX
KW Human; REG-like protein; RELP; immunoglobulin derived protein; Ig;
KW immunoglobulin; cytostatic; Ig agonist; immunoglobulin agonist; cancer;
KW protein therapy; RELP human Ig derived protein; chromosome 1p12-13.1;
KW PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200274916-A2.
XX
PD 26-SEP-2002.
XX
PF 14-MAR-2002; 2002WO-US07945.
XX
PR 16-MAR-2001; 2001US-276305P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heiskala M;
XX
DR WPI; 2003-103204/09.
XX
PT New isolated REG-like protein (RELP) human immunoglobulin derived
PT protein or specified portion or variant, useful for preventing or
PT treating a RELP protein mediated condition or malignant condition, e.g.
PT cancer -
XX
PS Example 6; Page 65; 101pp; English.
XX
CC The present invention describes a new isolated REG-like protein (RELP)
CC human immunoglobulin (Ig) derived protein. RELP comprises: (a) a human
CC variable and constant region; or (b) an isolated human Ig derived
CC protein or specified portion or variant encoded by a nucleic acid.
CC RELP has cytostatic activity and can be used as an Ig agonist and in
CC protein therapy. The RELP human Ig derived protein or a specified
CC portion or variant can be used for preventing or treating a RELP protein
CC mediated condition, malignant condition or disease condition, e.g.
CC cancer. The nucleic acids can be used in producing RELP Ig derived
CC protein. The human RELP protein of the present invention is located to

```
CC   chromosome 1p12-13.1. The present sequence represents a PCR primer for
CC   RELP, which is given in the exemplification of the present invention.
XX
SQ   Sequence 21 BP; 9 A; 7 C; 1 G; 4 T; 0 other;

  Query Match           100.0%;  Score 10;  DB 25;  Length 21;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         21 GAGTTTTGTT 12


RESULT 50
AAA59606/c
ID   AAA59606 standard; DNA; 23 BP.
XX
AC   AAA59606;
XX
DT   14-NOV-2000  (first entry)
XX
DE   PCR primer used to amplify DWF4 gene fragments.
XX
KW   DWF4; cytochrome P450 enzyme; brassinosteroid; 22alpha-hydroxylation;
KW   plant phenotype; cell elongation; PCR primer; ss.
XX
OS   Arabidopsis sp.
XX
PN   WO200047715-A2.
XX
PD   17-AUG-2000.
XX
PF   11-FEB-2000; 2000WO-US03820.
XX
PR   11-FEB-1999;   99US-0119657.
PR   11-FEB-1999;   99US-0119658.
XX
PA   (ARIZ-) ARIZONA BOARD OF REGENTS.
XX
PI   Azpiroz R,  Choe S,  Feldmann KA;
XX
DR   WPI; 2000-549142/50.
XX
PT   New isolated dwf4 polynucleotide useful for altering the phenotype of
PT   plants, for diagnostic assays and in the production of antibodies -
XX
PS   Example 2; Page 50; 113pp; English.
XX
CC   PCR primers AAA59600-12 were used to amplify fragments of a gene
CC   encoding a DWF4 polypeptide. The polypeptide is a cytochrome P450
CC   enzyme that mediates multiple steps in synthesis of brassinosteroids.
CC   Specifically, it mediates multiple 22alpha-hydroxylation steps in
CC   brassinosteroid biosynthesis. The DWF4 polynucleotide is used for
CC   altering the phenotype of a plant. DWF4 plants display a dramatic
CC   reduction in the length of different organs, and this size reduction
CC   is attributable to a defect in cell elongation. The DWF4 polynucleotides
CC   and polypeptides can be used in diagnostic assays and to generate
CC   antibodies, which can be used to produce immunogenic compositions.
XX
SQ   Sequence 23 BP; 8 A; 5 C; 3 G; 7 T; 0 other;

  Query Match           100.0%;  Score 10;  DB 21;  Length 23;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+04;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         20 GAGTTTTGTT 11


Search completed: January  2, 2004, 16:34:27
Job time : 293 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 2, 2004, 16:22:00 ; Search time 2603 Seconds
(without alignments)
93.371 Million cell updates/sec

Title: US-09-875-453B-5
Perfect score: 10
Sequence: 1 gagttttgtt 10

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 150 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		10	100.0	29	28	AZ851520	AZ851520 2M0153J18
c 2		10	100.0	31	28	AZ331527	AZ331527 1M0059F05
c 3		10	100.0	32	29	AL760995	AL760995 Arabidops
c 4		10	100.0	37	28	BH855860	BH855860 SALK_0844
c 5		10	100.0	46	9	AI663836	AI663836 uj06e03.x
6		10	100.0	46	28	AZ822694	AZ822694 2M0096H10
7		10	100.0	46	29	BZ290066	BZ290066 SALK_0234
c 8		10	100.0	47	29	BX131748	BX131748 Danio rer
9		10	100.0	51	9	AA529767	AA529767 vj12a02.r
c 10		10	100.0	52	10	BF643549	BF643549 NF026C12E
11		10	100.0	52	28	AZ767508	AZ767508 1M0566A17
c 12		10	100.0	60	28	BH863997	BH863997 SALK_0951
13		10	100.0	60	28	BH902064	BH902064 SALK_0912
14		10	100.0	61	9	AW552498	AW552498 L0213D06-
c 15		10	100.0	61	14	H67977	H67977 yr76a07.r1
c 16		10	100.0	62	14	N45259	N45259 yv26e03.r1
c 17		10	100.0	62	29	BZ767245	BZ767245 SALK_1385
c 18		10	100.0	65	28	AZ805922	AZ805922 2M0067N04
c 19		10	100.0	67	9	AI739301	AI739301 wi30b12.x
c 20		10	100.0	68	14	CD012715	CD012715 VVC033B04
c 21		10	100.0	69	29	CC178947	CC178947 SALK_0559
22		10	100.0	70	29	HSA247013	AJ247013 Homo sapi
c 23		10	100.0	72	28	AZ513784	AZ513784 1M0360H05
c 24		10	100.0	72	29	CC178946	CC178946 SALK_0559
25		10	100.0	73	10	BG508140	BG508140 sac98g04.
26		10	100.0	74	28	BH754805	BH754805 SALK_0441
c 27		10	100.0	76	10	BF590835	BF590835 7h42h01.x
28		10	100.0	76	28	AZ309821	AZ309821 1M0017C11
c 29		10	100.0	79	28	BH789918	BH789918 SALK_0528
30		10	100.0	82	9	AV533661	AV533661 AV533661
31		10	100.0	82	28	BH758559	BH758559 SALK_0280
32		10	100.0	82	28	BH790358	BH790358 SALK_0568
33		10	100.0	82	28	BH908137	BH908137 SALK_0458
34		10	100.0	83	12	BI980274	BI980274 ft74c04.x
c 35		10	100.0	84	29	CNS00Y3J	AL095677 Arabidops
c 36		10	100.0	85	29	CC457158	CC457158 SALK_1070
c 37		10	100.0	86	28	AZ825100	AZ825100 2M0100C08
c 38		10	100.0	86	29	BX175134	BX175134 Danio rer
c 39		10	100.0	87	10	BG058951	BG058951 nag51c01.
c 40		10	100.0	87	14	T11081	T11081 hbc628 Huma
c 41		10	100.0	87	28	AZ778481	AZ778481 2M0013K21
c 42		10	100.0	90	28	BH909222	BH909222 SALK_0524
c 43		10	100.0	90	29	DME546748	AJ546748 Drosophil
44		10	100.0	91	9	AW568695	AW568695 si60h08.y
45		10	100.0	91	29	AG227529	AG227529 Lotus jap
c 46		10	100.0	91	29	AG227530	AG227530 Lotus jap
c 47		10	100.0	91	29	AL751584	AL751584 Arabidops
c 48		10	100.0	93	28	AZ331545	AZ331545 1M0059J03
49		10	100.0	94	10	BE577421	BE577421 L48-2261T
c 50		10	100.0	95	28	BH212109	BH212109 SALK_0071
51		10	100.0	96	9	AA498709	AA498709 vh41c12.r
52		10	100.0	96	10	BE377756	BE377756 601229906
53		10	100.0	97	29	CC057326	CC057326 SALK_1189
c 54		10	100.0	98	29	BX292294	BX292294 Arabidops
c 55		10	100.0	99	28	AZ607111	AZ607111 1M0429G20
56		10	100.0	100	10	BE180679	BE180679 RC3-HT062
57		10	100.0	100	12	BI519601	BI519601 603061551
c 58		10	100.0	100	28	AZ454618	AZ454618 1M0256G10
c 59		10	100.0	101	10	BG658163	BG658163 TgESTzya4
c 60		10	100.0	101	12	BJ241595	BJ241595 BJ241595
61		10	100.0	101	28	AZ594915	AZ594915 1M0407K11
c 62		10	100.0	102	9	AU256871	AU256871 AU256871
c 63		10	100.0	102	9	AW362888	AW362888 PM2-CT029
c 64		10	100.0	102	13	BQ081420	BQ081420 san24c02.
c 65		10	100.0	102	13	BQ656791	BQ656791 HA02C20u
c 66		10	100.0	102	13	BQ741053	BQ741053 saq13g04.
c 67		10	100.0	102	28	AZ442208	AZ442208 1M0234A22
68		10	100.0	102	29	CC180333	CC180333 01S0415-0
69		10	100.0	103	9	AI327408	AI327408 mq39d05.x
c 70		10	100.0	103	28	AZ245182	AZ245182 RPCI-23-1
c 71		10	100.0	104	14	CD022852	CD022852 NXPV_086_
72		10	100.0	105	9	AW389427	AW389427 CM4-ST016
c 73		10	100.0	105	10	AW859019	AW859019 RC3-CT034
74		10	100.0	105	14	CA374648	CA374648 649118 NC
c 75		10	100.0	105	14	CA779223	CA779223 MPL384_8_
76		10	100.0	105	28	BH852704	BH852704 SALK_0754
77		10	100.0	106	28	BH219119	BH219119 1006084F0

	78	10	100.0	107	10	BE146428 BE146428 QV4-HT021
c	79	10	100.0	107	12	BI201660 BI201660 q4c11fs.f
	80	10	100.0	107	28	AZ596597 AZ596597 1M0409P22
	81	10	100.0	107	28	AQ481447 AQ481447 RPCI-11-2
	82	10	100.0	107	29	BZ371231 BZ371231 ie32f03.b
	83	10	100.0	108	28	AZ888039 AZ888039 RPCI-24-1
	84	10	100.0	109	9	AI503147 AI503147 vm99d05.x
	85	10	100.0	109	9	AI503835 AI503835 vm22a04.x
c	86	10	100.0	109	14	N53605 N53605 yz04b06.r1
c	87	10	100.0	109	28	AZ213650 AZ213650 Sheared D
	88	10	100.0	109	29	BX536223 BX536223 Arabidops
	89	10	100.0	109	29	CPA561374 AJ561374 Cryptospo
c	90	10	100.0	110	9	AV965756 AV965756 AV965756
c	91	10	100.0	110	10	BF942160 BF942160 nae87g09.
c	92	10	100.0	110	10	BF942207 BF942207 nae88g11.
	93	10	100.0	110	10	BF021759 BF021759 uy57g01.y
	94	10	100.0	110	14	CB006382 VVC033B04
c	95	10	100.0	110	14	CB178865 CB178865 in97f07.y
c	96	10	100.0	111	6	AL814079 A1814079 Triticum
	97	10	100.0	111	9	AA750933 AA750933 ISKH0261
	98	10	100.0	111	9	AW358248 AW358248 42265 MAR
c	99	10	100.0	111	13	BW134426 BW134426 BW134426
	100	10	100.0	111	28	BZ096185 BZ096185 CH230-142
	101	10	100.0	112	9	AJ280167 AJ280167 4A3A-AAL-
c	102	10	100.0	113	10	BF747482 BF747482 RC4-BT033
	103	10	100.0	113	12	BI041599 BI041599 PM0-NT031
c	104	10	100.0	113	12	BI201668 BI201668 q4d07fs.f
	105	10	100.0	113	13	BW165936 BW165936 BW165936
c	106	10	100.0	113	28	AQ527183 CITBI-E1-
c	107	10	100.0	114	9	AA591256 AA591256 vm22a04.r
c	108	10	100.0	114	10	BF361947 BF361947 QV2-NN004
	109	10	100.0	114	12	BJ036292 BJ036292
c	110	10	100.0	115	9	AA892725 AA892725 EST196528
c	111	10	100.0	115	10	BF361946 BF361946 QV2-NN004
c	112	10	100.0	115	10	BG139455 BG139455 EST479897
	113	10	100.0	115	10	BG462306 BG462306 947046G08
c	114	10	100.0	115	28	AZ621052 AZ621052 1M0454I06
	115	10	100.0	115	28	BH525450 BH525450 BOHLD38TF
	116	10	100.0	116	9	AW566029 AW566029 EST00021
c	117	10	100.0	116	10	BF746106 BF746106 RC4-BT033
c	118	10	100.0	116	12	BI054190 BI054190 PM3-GN037
c	119	10	100.0	116	13	BQ583033 BQ583033 E012099-0
	120	10	100.0	116	13	BQ844070 BQ844070 QGA12J01.
c	121	10	100.0	116	29	AL766963 AL766963 Arabidops
c	122	10	100.0	117	9	AA269255 AA269255 MA1NE009.
c	123	10	100.0	117	28	AQ026192 AQ026192 l(3)03884
	124	10	100.0	117	28	BH417926 BH417926 MuSite01-
	125	10	100.0	117	28	BH417932 BH417932 MuSite01-
	126	10	100.0	117	28	BH418018 BH418018 d1-2D3 Mu
	127	10	100.0	117	28	BH909412 BH909412 SALK_0534
c	128	10	100.0	118	14	CD023070 CD023070 NXPV_098_
c	129	10	100.0	118	14	D82586 D82586 HUMHBC2836_
c	130	10	100.0	118	14	T27431 T27431 hbc2836 Hum
c	131	10	100.0	118	28	BH336610 BH336610 CH230-50P
c	132	10	100.0	119	9	AI213541 AI213541 z3c02a1.r
c	133	10	100.0	119	13	BX085817 BX085817 BX085817
c	134	10	100.0	119	14	CB386098 CB386098 OSTF038H8
	135	10	100.0	119	28	BH846445 BH846445 SALK_0080
	136	10	100.0	119	29	CC101969 CC101969 CSU-K34.1
	137	10	100.0	120	9	AV186504 AV186504 AV186504
c	138	10	100.0	120	9	AA607325 AA607325 vm99d05.r
c	139	10	100.0	120	10	BF336401 BF336401 CM3-CT051
	140	10	100.0	120	13	BU032297 BU032297 QHJ20J17.
c	141	10	100.0	120	13	BW085116 BW085116 BW085116
c	142	10	100.0	120	14	CD012954 VVC047E09
	143	10	100.0	120	14	CD219837 CD219837 CCC1_59_G
	144	10	100.0	120	28	AZ931452 AZ931452 474.dhz84
c	145	10	100.0	121	9	AA512841 AA512841 ne85c07.s
	146	10	100.0	121	13	BQ164199 BQ164199 1091017D1
	147	10	100.0	121	28	BH212258 BH212258 SALK_0073
	148	10	100.0	122	9	AA087677 AA087677 mm27c09.r
	149	10	100.0	122	9	AI918059 AI918059 wc10a01.x
c	150	10	100.0	122	10	BF837698 BF837698 QV3-HT101

ALIGNMENTS

RESULT 1
AZ851520
LOCUS AZ851520 29 bp DNA linear GSS 21-FEB-2001
DEFINITION 2M0153J18R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0153J18 R, genomic survey sequence.
ACCESSION AZ851520
VERSION AZ851520.1 GI:13037599
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 29)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0153 row: J column: 18
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 29.
FEATURES Location/Qualifiers
source 1. .29
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0153J18"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 6 a 9 c 5 g 9 t
ORIGIN

Query Match 100.0%; Score 10; DB 28; Length 29;
Best Local Similarity 100.0%; Pred. No. 2.2e+05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         12 GAGTTTTGTT 21


RESULT 2
AZ331527/c
LOCUS       AZ331527                 31 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0059F05R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0059F05 R, genomic survey sequence.
ACCESSION   AZ331527
VERSION     AZ331527.1  GI:10394308
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0059  row: F  column: 05
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 31.
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0059F05"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AP129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       12 a      10 c       0 g       9 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         13 GAGTTTTGTT 4


RESULT 3
AL760995/c
LOCUS       AL760995                 32 bp    DNA     linear   GSS 18-JUN-2002
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-204F03-014508,
            genomic survey sequence.
ACCESSION   AL760995
VERSION     AL760995.1  GI:21501600
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 32)
  AUTHORS   Strizhov,N., Rosso,M., Li,Y. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion within the locus defined by clone K7M2. The
            sequences are generated at the MPI for Plant Breeding Research in
            the context of the GABI-Kat project. GABI-Kat is part of the German
            Plant Genomics program designated 'GABI'. Information on line
            availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .32
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-204F03-014508"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
BASE COUNT       13 a      9 c       3 g       7 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 29;  Length 32;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         15 GAGTTTTGTT 6
```

```
RESULT 4
BH855860/c
LOCUS       BH855860                37 bp    DNA     linear   GSS 08-JUL-2002
DEFINITION  SALK_084448.42.55.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_084448.42.55.x, genomic
            survey sequence.
ACCESSION   BH855860
VERSION     BH855860.1  GI:21705450
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 37)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_084448.42.55.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       18 a      8 c      4 g      7 t
ORIGIN

  Query Match            100.0%;  Score 10;  DB 28;  Length 37;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         16 GAGTTTTGTT 7


RESULT 5
AI663836/c
LOCUS       AI663836                  46 bp    mRNA    linear   EST 10-MAY-1999
DEFINITION  uj06e03.x1 Sugano mouse liver mlia Mus musculus cDNA clone
            IMAGE:1891132 3' similar to SW:ALBU_MERUN O35090 SERUM ALBUMIN
            PRECURSOR. ;, mRNA sequence.
ACCESSION   AI663836
VERSION     AI663836.1  GI:4767419
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished
COMMENT     Other_ESTs: uj06e03.y1
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:975456
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: custom primer used
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1891132"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Sugano mouse liver mlia"
                     /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a  DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert.  Size selection was
                     performed to exclude fragments <1.5kb.  Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science).  Custom primers  for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
BASE COUNT       13 a     12 c     11 g     10 t
ORIGIN

  Query Match            100.0%;  Score 10;  DB 9;  Length 46;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         27 GAGTTTTGTT 18


RESULT 6
AZ822694
LOCUS       AZ822694                  46 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0096H10F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0096H10 F, genomic survey sequence.
ACCESSION   AZ822694
VERSION     AZ822694.1  GI:12992602
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
```

```
            plasmid inserts
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0096  row: H  column: 10
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 46.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0096H10"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       14 a      4 c      9 g      19 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 46;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         14 GAGTTTTGTT 23


RESULT 7
BZ290066
LOCUS       BZ290066                 46 bp    DNA     linear   GSS 24-OCT-2002
DEFINITION  SALK_023483.34.55.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_023483.34.55.x, genomic
            survey sequence.
ACCESSION   BZ290066
VERSION     BZ290066.1  GI:24332810
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated exon of At4g30790.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_023483.34.55.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       12 a      11 c      9 g      14 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 29;  Length 46;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         15 GAGTTTTGTT 24


RESULT 8
BX131748/c
LOCUS       BX131748                 47 bp    DNA     linear   GSS 28-JAN-2003
DEFINITION  Danio rerio genomic clone DKEY-83F10, genomic survey sequence.
ACCESSION   BX131748
VERSION     BX131748.1  GI:27963018
KEYWORDS    GSS.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Humphray,S.J., Huckle,E. and Durham,J.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Unpublished
COMMENT     This sequence was generated from the T7 end of BAC 83F10. 83F10 is
            part of the Danio key BAC Library created by R. Plasterk and N.V.
            Keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-83F10"
                     /tissue_type="Testis"
                     /note="vector pIndigoBAC-536"
BASE COUNT       21 a      15 c      1 g      10 t
ORIGIN
```

```
  Query Match             100.0%;  Score 10;  DB 29;  Length 47;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         36 GAGTTTTGTT 27


RESULT 9
AA529767
LOCUS       AA529767                 51 bp    mRNA    linear   EST 22-JUL-1997
DEFINITION  vj12a02.r1 Barstead mouse proximal colon MPLRB6 Mus musculus cDNA
            clone IMAGE:921482 5', mRNA sequence.
ACCESSION   AA529767
VERSION     AA529767.1  GI:2272473
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:533698
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:921482"
                     /dev_stage="7 day juvenile"
                     /lab_host="DH10B"
                     /clone_lib="Barstead mouse proximal colon MPLRB6"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [AATTCGGATCCTTG], digested with Not I and cloned into the
                     Not I and Eco RI sites of the modified pT7T3 vector.
                     Library constructed by Bob Barstead. "
BASE COUNT       16 a     10 c      8 g     17 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 9;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         33 GAGTTTTGTT 42


RESULT 10
BF643549/c
LOCUS       BF643549                 52 bp    mRNA    linear   EST 20-DEC-2000
DEFINITION  NF026C12EC1F1098 Elicited cell culture Medicago truncatula cDNA
            clone NF026C12EC 5', mRNA sequence.
ACCESSION   BF643549
VERSION     BF643549.1  GI:11908770
KEYWORDS    EST.
SOURCE      Medicago truncatula (barrel medic)
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
            Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation -
            Center for Medicago Genomics Research
  JOURNAL   Unpublished
COMMENT     Contact: Dixon RA
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7302
            Fax: 580 221 7380
            Email: radixon@noble.org
            Insert Length: 52   Std Error: 0.00
            Plate: 026  row: C  column: 12
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Medicago truncatula"
                     /mol_type="mRNA"
                     /db_xref="taxon:3880"
                     /clone="NF026C12EC"
                     /tissue_type="Cell cultures derived from root tissues"
                     /dev_stage="Cell suspensions were subcultured every 14
                     days. Cells were induced six days after subculture"
                     /clone_lib="Elicited cell culture"
                     /note="Vector: Lambda Zap; Cells were induced with yeast
                     cell wall extracts equivalent to  50ug/ml glucose in the
                     final concentration. Samples were taken at  0.5, 1, 12 and
                     24 hours after induction. Equal amounts of RNA from  each
                     time point were pooled and used for mRNA isolation."
BASE COUNT       18 a     16 c      4 g     14 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 10;  Length 52;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         36 GAGTTTTGTT 27


RESULT 11
AZ767508
LOCUS       AZ767508                 52 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  1M0566A17R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0566A17 R, genomic survey sequence.
ACCESSION   AZ767508
VERSION     AZ767508.1  GI:12885672
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
```

```
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0566  row: A  column: 17
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 52.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0566A17"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       14 a      3 c      20 g      15 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 52;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         12 GAGTTTTGTT 21


RESULT 12
BH863997/c
LOCUS       BH863997                  60 bp    DNA     linear   GSS 05-AUG-2002
DEFINITION  SALK_095111 Arabidopsis thaliana TDNA insertion lines Arabidopsis
            thaliana genomic clone SALK_095111, genomic survey sequence.
ACCESSION   BH863997
VERSION     BH863997.1  GI:22099895
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_095111"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       22 a      11 c      10 g      17 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         44 GAGTTTTGTT 35


RESULT 13
BH902064
LOCUS       BH902064                  60 bp    DNA     linear   GSS 04-SEP-2002
DEFINITION  SALK_091210.33.70.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_091210.33.70.x, genomic
            survey sequence.
ACCESSION   BH902064
VERSION     BH902064.1  GI:22712945
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
```

```
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_091210.33.70.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       15 a      10 c      13 g      22 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db          6 GAGTTTTGTT 15


RESULT 14
AW552498
LOCUS       AW552498                 61 bp    mRNA    linear   EST 31-AUG-2000
DEFINITION  L0213D06-3 NIA Mouse Newborn Ovary cDNA Library Mus musculus cDNA
            clone L0213D06 3', mRNA sequence.
ACCESSION   AW552498
VERSION     AW552498.1  GI:7197921
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac
            ,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H.
            III, Becker,K.G. and Ko,M.S.H.
  TITLE     Genome-wide expression profiling of mid-gestation placenta and
            embryo using a 15,000 mouse developmental cDNA microarray
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
  MEDLINE   20381348
   PUBMED   10922068
COMMENT     Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: L0213  row: D  column: 06
            Seq primer: -21M13 Forward
            High quality sequence stop: 61
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .61
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:L0213D06-3"
                     /db_xref="taxon:10090"
                     /clone="L0213D06"
                     /sex="female"
                     /dev_stage="Newborn Ovary"
                     /lab_host="DH10B"
                     /clone_lib="NIA Mouse Newborn Ovary cDNA Library"
                     /note="Vector: pSPORT1 (Gibco/BRL Life Technology);
                     Site_1: SalI; Site_2: NotI; Total RNAs were extracted from
                     7 Newborn Ovary. The double-stranded  cDNA was synthesized
                     by Gibco's kit with an Oligo(dT) primer [NotI
                     primer-adapter from GibcoBRL]
                     [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT-3'] from
                     2.56ug of  total RNA . The double-stranded cDNAs were
                     treated with T4 DNA polymerase  and purified by
                     ethanol-precipitation. The cDNAs were ligated to
                     Lone-linker LL-Sal3 (include Sal1 sequence). The cDNAs
                     were purified by  phenol/chloroform and separated from
                     free linkers by Centricon 100. Then,  cDNAs were amplified
                     by long-range high fidelity PCR using Takara's Ex  Taq
                     polymerase. Then, the cDNAs were purified by
                     phenol/chloroform and  by Centricon 100. The cDNAs were
                     digested with SalI and NotI enzymes.  Then, the cDNAs were
                     size selected by Gibco's Size Fractionation Column.  The
                     cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid
                     vector. The  DH10B E. coli host was transformed with the
                     ligation mixture by chemical  method. The library was
                     constructed by Xiaohong Wang and Yulan Piao."
BASE COUNT       18 a       5 c      12 g      26 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 9;  Length 61;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         19 GAGTTTTGTT 28


RESULT 15
H67977/c
LOCUS       H67977                   61 bp    mRNA    linear   EST 18-OCT-1995
DEFINITION  yr76a07.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:211188 5' similar to SP:A42121 A42121 HELIX-LOOP-HELIX
            PROTEIN ;, mRNA sequence.
ACCESSION   H67977
VERSION     H67977.1  GI:1026717
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
   PUBMED   8889549
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 2780
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 2780   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .61
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3784029"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:211188"
                     /sex="male"
```

/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen 1NFLS"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 21 a 17 c 9 g 10 t 4 others
ORIGIN

Query Match 100.0%; Score 10; DB 14; Length 61;
Best Local Similarity 100.0%; Pred. No. 2.3e+05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 37 GAGTTTTGTT 28

RESULT 16
N45259/c
LOCUS N45259 62 bp mRNA linear EST 13-FEB-1996
DEFINITION yv26e03.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:243868 5', mRNA sequence.
ACCESSION N45259
VERSION N45259.1 GI:1186425
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 62)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Putative full length read
Seq primer: T7
High quality sequence stop: 273.
FEATURES Location/Qualifiers
source 1. .62
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3793001"
/db_xref="taxon:9606"
/clone="IMAGE:243868"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen 1NFLS"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 18 a 18 c 15 g 11 t
ORIGIN

Query Match 100.0%; Score 10; DB 14; Length 62;
Best Local Similarity 100.0%; Pred. No. 2.3e+05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 55 GAGTTTTGTT 46

RESULT 17
BZ767245/c
LOCUS BZ767245 62 bp DNA linear GSS 13-MAR-2003
DEFINITION SALK_138564.39.05.x Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_138564.39.05.x, genomic survey sequence.
ACCESSION BZ767245
VERSION BZ767245.1 GI:28939798
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 62)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
JOURNAL Unpublished
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 5' end of At5g22400.
Class: TDNA tagged.
FEATURES Location/Qualifiers
source 1. .62
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_138564.39.05.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT 20 a 7 c 10 g 25 t
ORIGIN

Query Match 100.0%; Score 10; DB 29; Length 62;
Best Local Similarity 100.0%; Pred. No. 2.3e+05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 31 GAGTTTTGTT 22

RESULT 18
AZ805922/c

```
LOCUS       AZ805922                65 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0067N04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0067N04 R, genomic survey sequence.
ACCESSION   AZ805922
VERSION     AZ805922.1  GI:12966733
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0067  row: N  column: 04
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 65.
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0067N04"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       23 a     17 c     13 g     12 t
ORIGIN

  Query Match           100.0%;  Score 10;  DB 28;  Length 65;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         22 GAGTTTTGTT 13


RESULT 19
AI739301/c
LOCUS       AI739301                67 bp    mRNA    linear   EST 18-JUN-1999
DEFINITION  wi30b12.x1 NCI_CGAP_Co16 Homo sapiens cDNA clone IMAGE:2391743 3'
            similar to TR:Q99523 Q99523 SORTILIN PRECURSOR. ;, mRNA sequence.
ACCESSION   AI739301
VERSION     AI739301.1  GI:5101282
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.
            , Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2391743"
                     /tissue_type="colon tumor, RER+"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Co16"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Co10 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1057416-1061255, and 1144584-1145351).
                     Subtraction by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT       14 a     25 c     13 g     15 t
ORIGIN

  Query Match           100.0%;  Score 10;  DB 9;  Length 67;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         25 GAGTTTTGTT 16


RESULT 20
CD012715/c
LOCUS       CD012715                68 bp    mRNA    linear   EST 02-MAY-2003
DEFINITION  VVC033B04_395039 An expressed sequence tag database for abiotic
            stressed berries of Vitis vinifera var. Chardonnay Vitis vinifera
            cDNA clone VVC033B04 3, mRNA sequence.
ACCESSION   CD012715
VERSION     CD012715.1  GI:30329453
KEYWORDS    EST.
SOURCE      Vitis vinifera
  ORGANISM  Vitis vinifera
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
```

```
            ; Vitaceae; Vitis.
REFERENCE   1  (bases 1 to 68)
  AUTHORS   Cushman,J.C.
  TITLE     An expressed sequence tag database for abiotic stressed berries of
            Vitis vinifera var. Chardonnay
  JOURNAL   Unpublished
COMMENT     Contact: Cushman JC
            Department of Biochemistry
            University of Nevada
            MS200, Reno, NV 89557-0014, USA
            Tel: 775-784-1918
            Fax: 775-784-1650
            Email: jcushman@unr.edu
            PCR PRimers
            FORWARD: T3 20mer
            BACKWARD: T7 21mer (backward)
            Plate: 033  row: B  column: 04
            Seq primer: T22V (V=A,C;G)
            High quality sequence stop: 68.
FEATURES             Location/Qualifiers
     source          1. .68
                     /organism="Vitis vinifera"
                     /mol_type="mRNA"
                     /db_xref="taxon:29760"
                     /clone="VVC033B04"
                     /tissue_type="berries"
                     /dev_stage="mixed; 8, 9, 11, 13, 15, 16 weeks daf"
                     /clone_lib="An expressed sequence tag database for abiotic
                     stressed berries of Vitis vinifera var. Chardonnay"
                     /note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
BASE COUNT       22 a     22 c     10 g     14 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 14;  Length 68;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         31 GAGTTTTGTT 22


RESULT 21
CC178947/c
LOCUS       CC178947                69 bp    DNA     linear   GSS 02-MAY-2003
DEFINITION  SALK_055936.38.90.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_055936.38.90.x, genomic
            survey sequence.
ACCESSION   CC178947
VERSION     CC178947.1  GI:30317498
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated intron of At1g78510.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_055936.38.90.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       20 a     21 c      7 g     21 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 29;  Length 69;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         21 GAGTTTTGTT 12


RESULT 22
HSA247013
LOCUS       HSA247013                70 bp    DNA     linear   GSS 24-JUN-1999
DEFINITION  Homo sapiens PAC trapped exon, clone 85M6 (70 bp), genomic survey
            sequence.
ACCESSION   AJ247013
VERSION     AJ247013.1  GI:5262870
KEYWORDS    GSS; PAC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Niederfuehr,A.
  JOURNAL   Thesis (1999) Universitaet Wuerzburg
REFERENCE   2  (bases 1 to 70)
  AUTHORS   Niederfuehr,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JUN-1999) Niederfuehr A., Physiologische Chemie I,
            Theodor-Boveri-Institut fuer Biowissenschaften, am Hubland, D-97074
            Wuerzburg, GERMANY
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="11"
                     /map="11p13"
                     /clone="85M6"
                     /clone_lib="RPCI PAC 1,3-5"
     exon            1. .70
                     /note="trapped"
BASE COUNT       19 a     14 c     16 g     21 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 29;  Length 70;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         16 GAGTTTTGTT 25


RESULT 23
AZ513784/c
```

```
LOCUS       AZ513784                72 bp    DNA     linear   GSS 05-OCT-2000
DEFINITION  1M0360H05F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0360H05 F, genomic survey sequence.
ACCESSION   AZ513784
VERSION     AZ513784.1  GI:10695100
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0360  row: H  column: 05
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 72.
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0360H05"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       27 a     11 c     14 g     20 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 72;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         69 GAGTTTTGTT 60


RESULT 24
CC178946/c
LOCUS       CC178946                72 bp    DNA     linear   GSS 02-MAY-2003
DEFINITION  SALK_055925.37.85.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_055925.37.85.x, genomic
            survey sequence.
ACCESSION   CC178946
VERSION     CC178946.1  GI:30317497
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated intron of At1g78510.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_055925.37.85.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       22 a     21 c      8 g     21 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 29;  Length 72;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         21 GAGTTTTGTT 12


RESULT 25
BG508140
LOCUS       BG508140                73 bp    mRNA    linear   EST 28-NOV-2001
DEFINITION  sac98g04.y1 Gm-c1073 Glycine max cDNA clone GENOME SYSTEMS CLONE
            ID: Gm-c1073-1088 5', mRNA sequence.
ACCESSION   BG508140
VERSION     BG508140.1  GI:13478797
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
```

```
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            High quality sequence stop: 72.
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1073-1088"
                     /tissue_type="seedlings induced for symptoms of SDS
                     (Sudden Death Syndrome) disease"
                     /dev_stage="2-3 weeks old"
                     /lab_host="DH10B"
                     /clone_lib="Gm-c1073"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; The cDNA library was constructed from mRNA isolated
                     from 2-3 week old seedlings that were induced for symptoms
                     of SDS (Sudden Death Syndrome) disease by the
                     translocation of culture filtrate of Fusarium solani f.
                     sp. glycines (Plant Cell Report 18:375-380). Cultivar
                     Williams 82 is  susceptible to the disease SDS. Plant
                     tissue (expanded leaves, folded leaves, and new shoots)
                     were collected at 1, 6, 24, and 48 hrs. after inoculation
                     and their mRNA pooled equally for cDNA construction. The
                     library was prepared using the Stratagene pBluescript II
                     SK(+) library construction kit. Complementary DNA was
                     synthesized from mRNA using a primer consisting of a
                     poly(dT) sequence with an XhoI restriction site. EcoRI
                     adaptors were ligated to the blunt-ended cDNA fragments
                     followed by XhoI digestion. The cDNA insert is protected
                     from XhoI digestion via methylation during first strand
                     synthesis. The cDNA fragments were directionally cloned
                     into the EcoRI-XhoI restriction site of the pBluescript
                     vector. The ligated cDNA fragments were transformed into
                     E.coli ElectroMax DH10B host cells. Plants were inoculated
                     by Shuxian Li (Glen Hartman lab, University of Illinois).
                     Library was constructed by Reena Philip and Steve Clough
                     (Lila Vodkin lab, University of Illinois)."
BASE COUNT       23 a      6 c     17 g     27 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 10;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         21 GAGTTTTGTT 30


RESULT 26
BH754805
LOCUS       BH754805                  74 bp    DNA     linear   GSS 01-MAR-2002
DEFINITION  SALK_044140.50.70.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_044140.50.70.x, genomic
            survey sequence.
ACCESSION   BH754805
VERSION     BH754805.1  GI:19035002
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 74)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .74
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_044140.50.70.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       24 a     12 c     14 g     24 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 74;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         46 GAGTTTTGTT 55


RESULT 27
BF590835/c
LOCUS       BF590835                  76 bp    mRNA    linear   EST 12-DEC-2000
DEFINITION  7h42h01.x1 NCI_CGAP_Co16 Homo sapiens cDNA clone IMAGE:3318673 3'
            similar to SW:TF1B_HUMAN Q13263 TRANSCRIPTION INTERMEDIARY FACTOR
            1-BETA ;, mRNA sequence.
ACCESSION   BF590835
VERSION     BF590835.1  GI:11683159
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.
            , Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
```

```
           Clone distribution: NCI-CGAP clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL, send email to:
           info@image.llnl.gov
           Trace considered overall poor quality
           Seq primer: -40UP from Gibco
           High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3318673"
                     /tissue_type="colon tumor, RER+"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Co16"
                     /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Co10 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1057416-1061255, and 1144584-1145351).
                     Subtraction by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT       20 a     21 c     21 g     14 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 10;  Length 76;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         40 GAGTTTTGTT 31


RESULT 28
AZ309821
LOCUS       AZ309821                 76 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0017C11F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0017C11 F, genomic survey sequence.
ACCESSION   AZ309821
VERSION     AZ309821.1  GI:10351196
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0017  row: C  column: 11
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 76.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0017C11"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        6 a     12 c     14 g     44 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 76;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         50 GAGTTTTGTT 59


RESULT 29
BH789918/c
LOCUS       BH789918                 79 bp    DNA     linear   GSS 02-APR-2002
DEFINITION  SALK_052802.43.55.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_052802.43.55.x, genomic
            survey sequence.
ACCESSION   BH789918
VERSION     BH789918.1  GI:19883016
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
```

```
                     /db_xref="taxon:3702"
                     /clone="SALK_052802.43.55.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       44 a      8 c     10 g     17 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 79;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         17 GAGTTTTGTT 8


RESULT 30
AV533661
LOCUS       AV533661                 82 bp    mRNA    linear   EST 06-SEP-2000
DEFINITION  AV533661 Arabidopsis thaliana flower buds Columbia Arabidopsis
            thaliana cDNA clone FB065h02F 3', mRNA sequence.
ACCESSION   AV533661
VERSION     AV533661.1  GI:8693944
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7, 175-180 (2000)
  MEDLINE   20363093
   PUBMED   10907847
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="FB065h02F"
                     /tissue_type="flower buds"
                     /clone_lib="Arabidopsis thaliana flower buds Columbia"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT       26 a     11 c     19 g     26 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 9;  Length 82;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         10 GAGTTTTGTT 19


RESULT 31
BH758559
LOCUS       BH758559                 82 bp    DNA     linear   GSS 01-MAR-2002
DEFINITION  SALK_028088.53.15.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_028088.53.15.x, genomic
            survey sequence.
ACCESSION   BH758559
VERSION     BH758559.1  GI:19044078
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_028088.53.15.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       19 a     13 c     17 g     33 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 82;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         65 GAGTTTTGTT 74


RESULT 32
BH790358
LOCUS       BH790358                 82 bp    DNA     linear   GSS 02-APR-2002
DEFINITION  SALK_056885.43.10.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_056885.43.10.x, genomic
            survey sequence.
ACCESSION   BH790358
VERSION     BH790358.1  GI:19883456
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
```

```
            Arabidopsis Genome
  JOURNAL   Unpublished
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_056885.43.10.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       30 a     14 c     13 g     25 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 82;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         56 GAGTTTTGTT 65


RESULT 33
BH908137
LOCUS       BH908137                 82 bp    DNA     linear   GSS 04-SEP-2002
DEFINITION  SALK_045875.41.40.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_045875.41.40.x, genomic
            survey sequence.
ACCESSION   BH908137
VERSION     BH908137.1  GI:22721070
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_045875.41.40.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       25 a     17 c      9 g     31 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 82;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         46 GAGTTTTGTT 55


RESULT 34
BI980274
LOCUS       BI980274                 83 bp    mRNA    linear   EST 24-OCT-2001
DEFINITION  ft74c04.x1 Gong zebrafish ovary Danio rerio cDNA clone
            IMAGE:5159023 3' similar to contains Alu repetitive element;, mRNA
            sequence.
ACCESSION   BI980274
VERSION     BI980274.1  GI:16367817
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 83)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            The library was constructed by Dr. Z. Gong. DNA Sequencing by:
            Washington University Genome Sequencing Center St. Louis. Please
            contact Zhiyuan Gong for further information on this library
            (National University of Singapore, Department of Biological
            Sciences, Lower Kent Ridge Road, Singapore 119260).
            Seq primer: T7 from Gibco.
FEATURES             Location/Qualifiers
     source          1. .83
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:5159023"
                     /sex="female"
                     /dev_stage="4-5 month"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Gong zebrafish ovary"
                     /note="Organ: ovary (pooled); Vector: pBluescript SK-;
                     Site_1: XhoI; Site_2: EcoRI; Poly A+ RNA was isolatd from
                     the ovaries of 2 female adult zebrafish (4-5 month  old).
                     cDNAs were made using oligo-dT primers and inserted into
                     lambda ZAP II vector  (Stratagene) by Dr. Z. Gong, in vivo
                     mass-excised to pBluescript SK- following the  Washington
```

University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Zhiyuan Gong for further information on
this library (National University of Singapore,
Department of Biological Sciences, Lower Kent Ridge Road,
Singapore 119260)."
BASE COUNT 10 a 7 c 14 g 52 t
ORIGIN

Query Match 100.0%; Score 10; DB 12; Length 83;
Best Local Similarity 100.0%; Pred. No. 2.4e+05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 34 GAGTTTTGTT 43

RESULT 35
CNS00Y3J/c
LOCUS CNS00Y3J 84 bp DNA linear GSS 29-JUN-1999
DEFINITION Arabidopsis thaliana genome survey sequence T7 end of BAC T15I13 of TAMU library from strain Columbia of Arabidopsis thaliana, genomic survey sequence.
ACCESSION AL095677
VERSION AL095677.1 GI:5303832
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 84)
AUTHORS Salanoubat,M., Choisne,N., Artiguenave,F., Brottier,P., Wincker,P., Samson,D., Saurin,W., Weissenbach,J. and Quetier,F.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 84)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
FEATURES Location/Qualifiers
source 1. .84
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="T15I13"
/clone_lib="TAMU"
/note="end : T7"
BASE COUNT 37 a 14 c 12 g 21 t
ORIGIN

Query Match 100.0%; Score 10; DB 29; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.4e+05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 61 GAGTTTTGTT 52

RESULT 36
CC457158/c
LOCUS CC457158 85 bp DNA linear GSS 30-MAY-2003
DEFINITION SALK_107080.42.10.x Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_107080.42.10.x, genomic survey sequence.
ACCESSION CC457158
VERSION CC457158.1 GI:31218734
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 85)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
JOURNAL Unpublished
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 3' end of At3g59290.
Class: TDNA tagged.
FEATURES Location/Qualifiers
source 1. .85
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_107080.42.10.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT 25 a 18 c 12 g 30 t
ORIGIN

Query Match 100.0%; Score 10; DB 29; Length 85;
Best Local Similarity 100.0%; Pred. No. 2.4e+05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTTGTT 10
||||||||||
Db 52 GAGTTTTGTT 43

RESULT 37
AZ825100/c
LOCUS AZ825100 86 bp DNA linear GSS 20-FEB-2001
DEFINITION 2M0100C08F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0100C08 F, genomic survey sequence.
ACCESSION AZ825100
VERSION AZ825100.1 GI:12995008
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 86)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

```
           Tel: 801 585 5606
           Fax: 801 585 7177
           Email: ddunn@genetics.utah.edu
           Insert Length: 10000   Std Error: 0.00
           Plate: 0100  row: C  column: 08
           Seq primer: CGTTGTAAAACGACGGCCAGT
           Class: plasmid ends
           High quality sequence stop: 86.
FEATURES             Location/Qualifiers
     source          1. .86
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0100C08"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       28 a     20 c     14 g     24 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 86;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         49 GAGTTTTGTT 40


RESULT 38
BX175134/c
LOCUS       BX175134                 86 bp    DNA     linear   GSS 13-MAR-2003
DEFINITION  Danio rerio genomic clone DKEY-184L8, genomic survey sequence.
ACCESSION   BX175134
VERSION     BX175134.1  GI:28006844
KEYWORDS    GSS.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 86)
  AUTHORS   Humphray,S.J., Huckle,E. and Durham,J.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Unpublished
COMMENT     This sequence was generated from the SP6 end of BAC 184L8. 184L8 is
            part of the Daniokey BAC Library created by R. Plasterk and N.V.
            Keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES             Location/Qualifiers
     source          1. .86
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-184L8"
                     /tissue_type="Testis"
                     /note="vector pIndigoBAC-536"
BASE COUNT       53 a     10 c     16 g      7 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 29;  Length 86;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         34 GAGTTTTGTT 25


RESULT 39
BG058951/c
LOCUS       BG058951                 87 bp    mRNA    linear   EST 25-JAN-2001
DEFINITION  nag51c01.y1 NCI_CGAP_Co27 Homo sapiens cDNA clone IMAGE:4205160 5',
            mRNA sequence.
ACCESSION   BG058951
VERSION     BG058951.1  GI:12525948
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 87)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .87
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4205160"
                     /tissue_type="adenocarcinoma (mucinous component)"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Co27"
                     /note="Organ: colon; Vector: pAMP1; mRNA made from colonic
                     adenocarcinoma, cDNA made by oligo-dT priming.
                     Directionally cloned into UDG sites. Size-selected on
                     agarose gel,  average insert size 300 bp. Primary library.
                     cDNA Library Preparation:  David B. Krizman, Ph.D.
                     Reference: Krizman et al. (1996) Cancer Research
                     56:5380-5383."
BASE COUNT       41 a     10 c     17 g     19 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 10;  Length 87;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         37 GAGTTTTGTT 28


RESULT 40
```

```
T11081/c
LOCUS       T11081                  87 bp    mRNA    linear   EST 29-NOV-1993
DEFINITION  hbc628 Human pancreatic islet Homo sapiens cDNA clone hbc628 5'end
            similar to pancreatic lipase, mRNA sequence.
ACCESSION   T11081
VERSION     T11081.1  GI:391235
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 87)
  AUTHORS   Takeda,J., Yano,H., Eng,S., Zeng,Y. and Bell,G.I.
  TITLE     A molecular inventory of human pancreatic islets: sequence analysis
            of 1000 cDNA clones
  JOURNAL   Hum. Mol. Genet. 2, 1793-1798 (1993)
  MEDLINE   94108427
   PUBMED   7506601
COMMENT     Contact: Bell GI or Takeda J
            HHMI
            Univ. of Chicago
            5841 S. Maryland Ave.,MC1028,Chicago IL 60637
            Tel: 3127029116
            Fax: 3127020271
            Email: g-bell@uchicago.edu
            Seq primer: SK primer.
FEATURES             Location/Qualifiers
     source          1. .87
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="hbc628"
                     /clone_lib="Human pancreatic islet"
                     /note="Vector: Lambda ZAPII; Site_1: Eco RI; Site_2: Xho
                     I; mRNA was prepared from normal adult human islets. cDNA
                     was directionally synthesized from the Xho I in the vector
                     to the EcoRI site. cDNA was size fractionated to remove
                     sequences <1000 bp in size."
BASE COUNT       17 a     26 c     22 g     22 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 14;  Length 87;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         76 GAGTTTTGTT 67


RESULT 41
AZ778481/c
LOCUS       AZ778481                87 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0013K21R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0013K21 R, genomic survey sequence.
ACCESSION   AZ778481
VERSION     AZ778481.1  GI:12908169
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 87)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0013  row: K  column: 21
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 87.
FEATURES             Location/Qualifiers
     source          1. .87
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0013K21"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       21 a     22 c     21 g     23 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 87;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         55 GAGTTTTGTT 46


RESULT 42
BH909222/c
LOCUS       BH909222                90 bp    DNA     linear   GSS 04-SEP-2002
DEFINITION  SALK_052462.47.95.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_052462.47.95.x, genomic
            survey sequence.
ACCESSION   BH909222
VERSION     BH909222.1  GI:22722155
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 90)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished
COMMENT     Contact: Joseph R. Ecker
```

```
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES              Location/Qualifiers
     source           1. .90
                      /organism="Arabidopsis thaliana"
                      /mol_type="genomic DNA"
                      /strain="Columbia 0"
                      /db_xref="taxon:3702"
                      /clone="SALK_052462.47.95.x"
                      /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                      /note="PCR was performed on Arabidopsis thaliana lines
                      each of which contains  one or more TDNA insertion
                      elements. The resultant fragment for each  line was
                      directly sequenced to determine the genomic sequence at
                      the  site of insertion. Details of the protocols used can
                      be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       35 a     20 c     16 g     19 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 90;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         59 GAGTTTTGTT 50


RESULT 43
DME546748/c
LOCUS       DME546748                 90 bp    DNA     linear   GSS 24-FEB-2003
DEFINITION  Drosophila melanogaster flanking sequence of RS P element insertion
            P{RS5}5-SZ-3976, clone library P{RS5}, genomic survey sequence.
ACCESSION   AJ546748
VERSION     AJ546748.1  GI:28554883
KEYWORDS    GSS; genome survey sequence.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1
  AUTHORS   Ryder,E.J., Ashburner,M., Bagunya,J., Blows,F., Bucheton,A.,
            Coulson,D., Dickson,B., Drummond,J., Glover,D., Gunton,N.,
            Hafen,E., Hall,S., Heisenberg,M., Lepesant,J.A., Maroy,P.,
            Mechler,B., O'Kane,C., Pflugfelder,G., Rasmuson-Lestander,A.,
            Reuter,G., Roote,J., Szidonya,J., Wang,S., Webster,J. and
            Russell,S.
  TITLE     Mapping of RS P element insertions in Drosophila melanogaster for
            the DrosDel second generation deficiency kit
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 90)
  AUTHORS   Ryder,E.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-FEB-2003) Ryder E.J., Department of Genetics,
            University of Cambridge, Downing Street, CB2 3EH, UNITED KINGDOM
COMMENT     The insertion point of the P element is before base 1 of the
            sequence. Further information about this P element insertion line
            can be found at http://www.flyseq.org.uk and
            http://www.drosdel.org.uk.
FEATURES              Location/Qualifiers
     source           1. .90
                      /organism="Drosophila melanogaster"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:7227"
                      /chromosome="2R"
                      /clone="P{RS5}5-SZ-3976"
                      /clone_lib="P{RS5}"
                      /note="read=5' end"
     misc_feature     1. .90
                      /note="P element insertion in the 3' to 5' orientation"
BASE COUNT       28 a     23 c     15 g     24 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 29;  Length 90;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         47 GAGTTTTGTT 38


RESULT 44
AW568695
LOCUS       AW568695                  91 bp    mRNA    linear   EST 03-DEC-2001
DEFINITION  si60h08.y1 Gm-r1030 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-r1030-3352 5', mRNA sequence.
ACCESSION   AW568695
VERSION     AW568695.1  GI:7233348
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 91)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com.
FEATURES              Location/Qualifiers
     source           1. .91
                      /organism="Glycine max"
                      /mol_type="mRNA"
                      /db_xref="taxon:3847"
                      /clone="GENOME SYSTEMS CLONE ID: Gm-r1030-3352"
                      /lab_host="DH10B"
                      /clone_lib="Gm-r1030"
                      /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                      cDNA library was constructed from mRNA isolated from
                      immature cotyledons  of greenhouse grown plants
                      (individual seed fresh weight of 100-300mg).  The  library
                      was prepared using the Life Technologies pSuperScript cDNA
                      library  construction kit.  Complementary DNA was
                      synthesized from mRNA using a poly(dT)  sequence with a
                      NotI restriction site.  SalI linkers adapters were ligated
                      to  the blunt-ended cDNA fragments followed by NotI
                      digestion.  The cDNA fragments  were directionally cloned
                      into the NotI-SalI restriction site of the pSPORT1
                      vector.  The ligated cDNA fragments were transformed into
                      E. coli ElectroMax  DH10B host cells.  This library was
                      constructed by Dr. Lila Vodkin and  Dr. Anu Khanna.  Note
                      that Gm-r1030 is a re-rack of Gm-c1007."
```

```
BASE COUNT       22 a      9 c     17 g     43 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 9;  Length 91;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         78 GAGTTTTGTT 87


RESULT 45
AG227529
LOCUS       AG227529                 91 bp    DNA     linear   GSS 12-DEC-2002
DEFINITION  Lotus japonicus DNA, clone:LjB141g19_f, genomic survey sequence.
ACCESSION   AG227529
VERSION     AG227529.1  GI:26538153
KEYWORDS    GSS.
SOURCE      Lotus japonicus
  ORGANISM  Lotus japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1
  AUTHORS   Sato,S., Nakamura,Y. and Tabata,S.
  TITLE     Lotus japonicus BAC End sequences
  JOURNAL   Published Only in Database (2002)
REFERENCE   2  (bases 1 to 91)
  AUTHORS   Sato,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-2002) Shusei Sato, Kazusa DNA Research Institute,
            The First Laboratory for Plant Gene Research; 2-6-7
            Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:ssato@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/,
            Tel:81-438-52-3935(ex.2336), Fax:81-438-52-3934)
FEATURES             Location/Qualifiers
     source          1. .91
                     /organism="Lotus japonicus"
                     /mol_type="genomic DNA"
                     /strain="Miyakojima MG-20"
                     /db_xref="taxon:34305"
                     /clone="LjB141g19_f"
                     /clone_lib="genomic BAC library"
                     /note="VECTOR:pBeloBAC11"
BASE COUNT       22 a      19 c      9 g      41 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 29;  Length 91;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         53 GAGTTTTGTT 62


RESULT 46
AG227530/c
LOCUS       AG227530                 91 bp    DNA     linear   GSS 12-DEC-2002
DEFINITION  Lotus japonicus DNA, clone:LjB141g19_r, genomic survey sequence.
ACCESSION   AG227530
VERSION     AG227530.1  GI:26538154
KEYWORDS    GSS.
SOURCE      Lotus japonicus
  ORGANISM  Lotus japonicus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
            Lotus.
REFERENCE   1
  AUTHORS   Sato,S., Nakamura,Y. and Tabata,S.
  TITLE     Lotus japonicus BAC End sequences
  JOURNAL   Published Only in Database (2002)
REFERENCE   2  (bases 1 to 91)
  AUTHORS   Sato,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-2002) Shusei Sato, Kazusa DNA Research Institute,
            The First Laboratory for Plant Gene Research; 2-6-7
            Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:ssato@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/,
            Tel:81-438-52-3935(ex.2336), Fax:81-438-52-3934)
FEATURES             Location/Qualifiers
     source          1. .91
                     /organism="Lotus japonicus"
                     /mol_type="genomic DNA"
                     /strain="Miyakojima MG-20"
                     /db_xref="taxon:34305"
                     /clone="LjB141g19_r"
                     /clone_lib="genomic BAC library"
                     /note="VECTOR:pBeloBAC11"
BASE COUNT       38 a      9 c      22 g     22 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 29;  Length 91;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         39 GAGTTTTGTT 30


RESULT 47
AL751584/c
LOCUS       AL751584                 91 bp    DNA     linear   GSS 17-JUN-2002
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-005D06-014769,
            genomic survey sequence.
ACCESSION   AL751584
VERSION     AL751584.1  GI:21484081
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 91)
  AUTHORS   Rosso,M., Strizhov,N., Li,Y. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion within the locus defined by clone f15b18.
            The sequences are generated at the MPI for Plant Breeding Research
            in the context of the GABI-Kat project. GABI-Kat is part of the
            German Plant Genomics program designated 'GABI'. Information on
            line availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .91
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
```

```
                     /db_xref="taxon:3702"
                     /clone="GK-005D06-014769"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC106. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
BASE COUNT       34 a     21 c     15 g     21 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 29;  Length 91;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         37 GAGTTTTGTT 28


RESULT 48
AZ331545/c
LOCUS       AZ331545                  93 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0059J03R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0059J03 R, genomic survey sequence.
ACCESSION   AZ331545
VERSION     AZ331545.1  GI:10394343
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0059  row: J  column: 03
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 93.
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0059J03"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       45 a     20 c      7 g     21 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 93;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         19 GAGTTTTGTT 10


RESULT 49
BE577421
LOCUS       BE577421                  94 bp    mRNA    linear   EST 20-FEB-2001
DEFINITION  L48-2261T3 Ice plant Lambda Uni-Zap XR expression library, 48 hours
            NaCl treatment Mesembryanthemum crystallinum cDNA clone L48-2261
            5', mRNA sequence.
ACCESSION   BE577421
VERSION     BE577421.1  GI:9827220
KEYWORDS    EST.
SOURCE      Mesembryanthemum crystallinum (common iceplant)
  ORGANISM  Mesembryanthemum crystallinum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Caryophyllales; Aizoaceae; Mesembryanthemum.
REFERENCE   1  (bases 1 to 94)
  AUTHORS   Cushman, J.C.
  TITLE     An expressed sequence tag database for the common ice plant,
            Mesembryanthemum crystallinum
  JOURNAL   Unpublished
COMMENT     Contact: Cushman JC
            Department of Biochemistry
            University of Nevada
            MS200, Reno, NV 89557-0014, USA
            Tel: 775-784-1918
            Fax: 775-784-1650
            Email: jcushman@unr.edu
            PCR PRimers
            FORWARD: T7
            BACKWARD: T3
            Plate: L48-23  row: F  column: 1
            Seq primer: T3
            High quality sequence stop: 94
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .94
                     /organism="Mesembryanthemum crystallinum"
                     /mol_type="mRNA"
                     /db_xref="taxon:3544"
                     /clone="L48-2261"
                     /tissue_type="Leaf, 48 h 0.4M NaCl"
                     /dev_stage="Six week old"
                     /clone_lib="Ice plant Lambda Uni-Zap XR expression library
                     , 48 hours NaCl treatment"
                     /note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI"
BASE COUNT       30 a     18 c     20 g     26 t
ORIGIN
```

```
  Query Match             100.0%;  Score 10;  DB 10;  Length 94;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         57 GAGTTTTGTT 66


RESULT 50
BH212109/c
LOCUS       BH212109                 95 bp    DNA     linear   GSS 24-OCT-2001
DEFINITION  SALK_007106 Arabidopsis thaliana TDNA insertion lines Arabidopsis
            thaliana genomic clone SALK_007106, genomic survey sequence.
ACCESSION   BH212109
VERSION     BH212109.1  GI:16393007
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 95)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within 300 bases of the 5' end of
            At3g46500.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .95
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_007106"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       32 a     22 c     28 g     13 t
ORIGIN

  Query Match             100.0%;  Score 10;  DB 28;  Length 95;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAGTTTTGTT 10
              ||||||||||
Db         40 GAGTTTTGTT 31


Search completed: January  2, 2004, 18:03:52
Job time : 2627 secs
```